ÉPREUVE COMMUNE DE TIPE - Partie D

TITRE :

LA DÉMOCRATIE, OBJET D'ÉTUDE MATHÉMATIQUE

Temps de préparation : ....2 h 15 minutes
Temps de présentation devant les examinateurs : ....10 minutes
Entretien avec les examinateurs : ....10 minutes

GUIDE POUR LES CANDIDAT(E)S :
Le dossier ci-joint comporte au total : 20 pages (dont celle-ci).
Le document consiste de deux parties, chacune est relativement indépendante de l'autre.
Ces deux documents sont issus du site « Images des mathématiques » et ont été écrits par Rémi Peyre.

TRAVAIL SUGGÉRÉ **:**

Lire les deux documents rapidement, sans s'attarder sur les démonstrations.
Effectuer une synthèse en s'attachant à montrer un ou plusieurs problèmes provenant de la confrontation entre deux points de vue sur la démocratie.
On s'efforcera de présenter au moins une démonstration en dégageant les points clefs, mais en réservant les détails pour un éventuel complément lors des questions.
On cherchera surtout à bien expliquer la nature des problèmes en considérant des exemples non issus directement du texte.

CONSEILS GENERAUX POUR LA PREPARATION DE L'EPREUVE :

* Lisez le dossier en entier dans un temps raisonnable (moins d'une heure).
* Réservez du temps pour préparer l'exposé (au moins 20 minutes).
* Préparez une introduction <u>et</u> une conclusion.
* Ne passez pas trop de temps sur une démonstration durant l'exposé.
* Ne surchargez pas vos transparents de mathématiques : restez lisible.
* Aérez et soignez vos transparents.

- En fin de préparation, rassemblez et ordonnez soigneusement TOUS les documents (transparents, etc.) dont vous comptez vous servir pendant l'oral, ainsi que le dossier, les transparents et les brouillons utilisés pendant la préparation. En entrant dans la salle d'oral, vous devez être prêts à débuter votre exposé.

<u>Le jour de l'oral, les consignes suivantes seront également présentes :</u>
- Vous pouvez écrire sur le présent dossier, le surligner, le découper … mais tout sera à remettre aux examinateurs en fin d'oral.
- A la fin de l'oral, vous devez remettre aux examinateurs le présent dossier, les transparents et les brouillons utilisés pour cette partie de l'oral, ainsi que TOUS les transparents et autres documents présentés pendant votre prestation.

Les mathématiques de la démocratie, I

La démocratie, objet d’étude mathématique

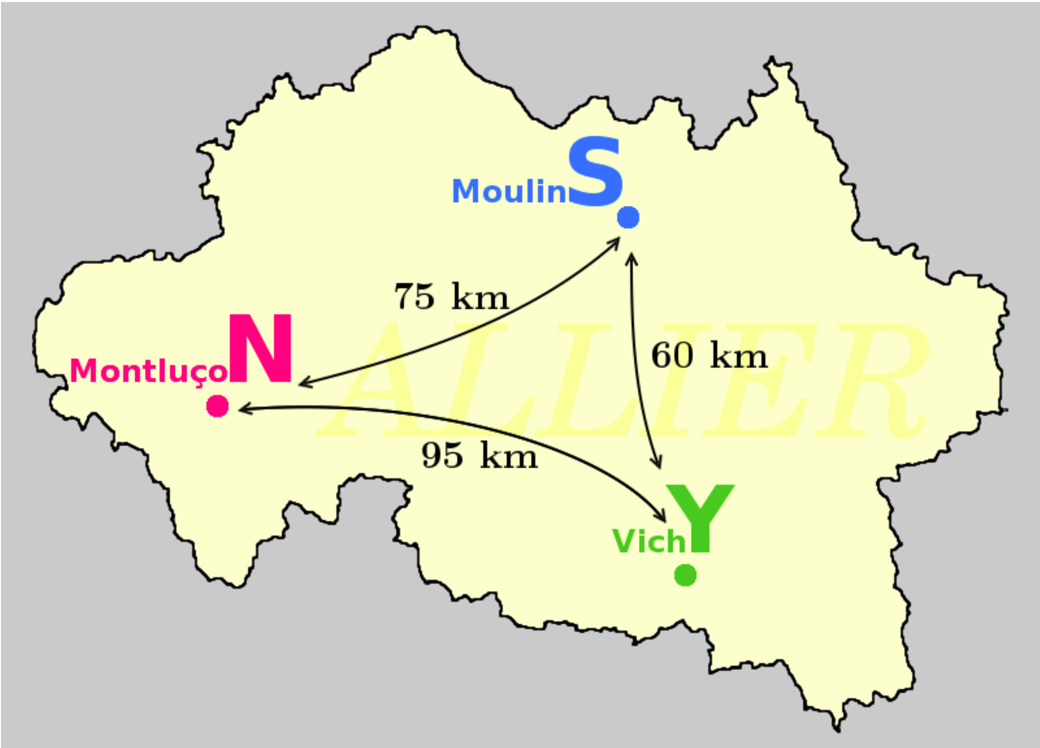
Le 16 avril 2012, par **Rémi Peyre**
Maître de conférences à l'École des Mines de Nancy. (**page web**)

Une démocratie est un régime politique dans lequel les décisions sont prises en fonction de la volonté du peuple. Mais qu’est-ce que la volonté du peuple ? En d’autres termes, quelle est la bonne façon de tenir compte des préférences de chaque individu pour en déduire la préférence collective ? Depuis quelques décennies, les mathématiciens se sont penchés sur ces questions, et sont arrivés à des conclusions... surprenantes ! (Ce texte est le premier d’une série de trois articles sur les problématiques mathématiques soulevées par l’étude du concept de démocratie).

Introduction : La démocratie ? *Quelle* démocratie ?

Le dilemme des Alliérins

Suite à une révolution, le département de l’Allier est devenu indépendant ! La Constitution du nouvel État prévoit que des élections démocratiques soient organisées pour désigner sa capitale. Il se trouve que l’Allier présente la particularité d’avoir trois villes principales de tailles comparables, villes que nous noterons par leur dernière lettre : Montluçon (N), Moulins (S) et Vichy (Y). Chacune des ces villes a déposé une candidature pour devenir capitale, et chaque citoyen souhaite naturellement que la future capitale soit la plus proche possible de chez lui. Ainsi [voir la carte ci-dessous], les habitants de N, à défaut de voir leur ville choisie, préfèreraient S à Y, les habitants de S préfèreraient Y à N, et les habitants de Y préfèreraient S à N. (Nous négligerons ici les citoyens n’habitant aucune des trois villes).



L’Allier

Carte de l’Allier montrant les trois villes et les distances les séparant.

Si la Constitution dit bien que c’est par une élection démocratique que doit être désignée la capitale, elle ne précise en revanche pas quelle méthode électorale doit être suivie. Plusieurs méthodes concurrentes, toutes clairement démocratiques, ont été proposées :

- Méthode A : Chaque électeur vote pour une ville ; la ville qui reçoit le plus de suffrages gagne.
- Méthode B : Dans un premier tour, chaque électeur vote pour une ville, puis un second tour est organisé entre les deux villes ayant reçu le plus de suffrages au premier tour ; la ville qui reçoit le plus de suffrages au second tour gagne.

- Méthode C : Chaque électeur vote *contre* une ville ; la ville qui reçoit le *moins* de suffrages gagne.
- Méthode D : Dans un premier tour, chaque électeur vote contre une ville, puis un second tour est organisé entre les deux villes ayant reçu le moins de suffrages au premier tour ; la ville qui reçoit le moins de suffrages au second tour gagne.
- Méthode E : Chaque électeur classe les villes par ordre de préférence ; sa ville préférée marque 2 points, la suivante 1 point et la dernière aucun. La ville qui obtient le meilleur total gagne.

Diantre, que de possibilités ! Mais bon, vu que ces méthodes sont toutes démocratiques, peut-être donnent-elles en fait la même vainqueuse... Hélas, non. Imaginons par exemple que 40 % des électeurs habitent N, 35 % habitent Y et 25 % habitent S. Alors :

- Suivant la méthode A, N gagne par 40 % des suffrages contre 35 % pour Y et 25 % pour S.
- Suivant la méthode B, Y gagne au second tour contre N par 60 % des suffrages contre 40 %, vu que les habitants de S se reportent sur Y au second tour.
- Suivant la méthode C, S gagne par 0 % des suffrages contre 40 % pour Y et 60 % pour N, vu que les habitants de S et de Y votent contre N et que ceux de N votent contre Y.
- Suivant la méthode D, S gagne au second tour contre Y par 35 % des suffrages contre 65 %, puisqu'après l'élimination de N les habitants de S votent contre Y et ceux de Y contre S.
- Suivant la méthode E, S gagne avec une moyenne de 1,25 points par électeur (2 points de la part des habitants de S, et 1 point de ceux de N et de Y), suivi de Y avec 0,95 points (2 points de la part des habitants de Y, 1 point de ceux de S, et aucun de ceux de N), et de N avec 0,80 points (2 points de la part des habitants de N, et aucun de ceux de Y et de S).

Notez qu'ici les méthodes C, D et E donnent la même vainqueuse, mais qu'elles auraient pu aboutir à des vainqueuses contradictoires pour une autre répartition des électeurs entre les trois villes... Le tableau ci-dessous montre ainsi quelles seraient les vainqueuses suivant les différentes méthodes pour diverses répartitions des électeurs (le cas que nous venons de détailler est la ligne sur fond gris).

Les vainqueuses donnés par les différentes méthodes pour diverses répartitions des électeurs.

Proportion d'habitants de N	Habitants de S	Habitants de Y	Vainqueuse selon la méthode A	Selon B	Selon C	Selon D	Selon E
60 %	15 %	25 %	N	N	S	N	N
60 %	25 %	15 %	N	N	S	N	S
40 %	25 %	35 %	**N**	**Y**	**S**	**S**	**S**
40 %	35 %	25 %	N	S	S	S	S
30 %	40 %	30 %	S	S	S	S	S
25 %	35 %	40 %	Y	S	S	S	S
35 %	25 %	40 %	Y	Y	S	S	S
20 %	20 %	60 %	Y	Y	S	Y	Y

La conclusion qu'on tire de cet exemple introductif, c'est que le concept de "volonté du peuple" se révèle ambigu, dans la mesure où plusieurs définitions a priori raisonnables de cette volonté sont en fait contradictoires... Par conséquent, le concept de démocratie, en tant que régime politique dans lequel les décisions sont prises en fonction de la volonté du peuple, est également ambigu ! Ainsi, derrière l'idée "naturelle" de démocratie, se pose le délicat problème suivant : comment déterminer la volonté du peuple à partir des préférences individuelles ? Y a-t-il des méthodes qui soient meilleures que d'autres, et si oui lesquelles ?

Or, dès lors qu'il s'agit de rendre rigoureuse l'étude d'un concept, on entre dans le terrain de jeu des mathématiciens... Il n'est donc pas surprenant que depuis la seconde moitié du XXe siècle, des dizaines de chercheurs se soient penchés sur les problèmes afférents à la théorie des élections démocratiques [1]. Leurs conclusions, nous le verrons, sont parfois très surprenantes...!

Cadre de cet article

L'objet de cet article (et de ceux à suivre) est d'analyser les différentes méthodes électorales démocratiques. Il y a différents contextes dans lesquels on peut vouloir recourir à un scrutin : choisir un président, composer une assemblée de représentants, ajuster les composantes d'un budget... Ces différents contextes peuvent conduire à des problématiques mathématiques différentes. Dans cette série d'articles, nous allons nous limiter à un cadre précis, mais malgré tout très général :

- On s'intéresse aux situations où il s'agit de choisir une seule option parmi plusieurs. On peut en fait toujours se ramener à ce cadre : ainsi, choisir une entrée, puis un plat, puis un dessert est équivalent à choisir directement un menu ; de même, classer les candidats d'un concours équivaut à choisir un classement parmi tous ceux possibles...
- On demande que l'élection se déroule en un seul tour. Là non plus ce n'est pas une vraie restriction, attendu qu'en revanche les bulletins de vote peuvent contenir des informations complexes. Une élection selon la méthode B ci-dessus, par exemple, peut très bien être organisée en un seul scrutin : dans cette organisation, chaque bulletin comportera, d'une part le choix de l'électeur au premier tour, d'autre part l'ensemble de ses choix au second tour pour tous les seconds tours possibles, ce qui donne bien une méthode équivalente au protocole à deux tours.
- Les options proposées sont en nombre fini (mais peut-être très grand), de même que les types de suffrages possibles [2]. Il s'agit là d'une exigence essentiellement technique destinée à éviter certaines difficultés mathématiques ; en pratique ce n'est pas une restriction très importante.

Les hypothèses

Puisque nous parlons ici de mathématiques, il nous faut aussi donner un sens précis à ce que nous entendons par « méthode électorale démocratique ». Dans cette série d'articles, nous imposerons les hypothèses suivantes :

- Tous les électeurs ont le même statut : aucun électeur n'est privilégié par rapport aux autres [3], et il n'y a pas non plus de système de "circonscriptions" qui établirait des relations particulières entre certains d'entre eux [4].
- De même, toutes les options ont le même statut. En particulier, il n'y a aucune restriction sur les préférences que peut avoir un électeur : dans une

élection réelle, cela signifierait par exemple qu'un électeur peut très bien avoir en premier choix le candidat "de droite", en second choix le candidat "de gauche" et en troisième choix le candidat "du centre", puisqu'ici les mots « gauche », « centre » et « droite » n'ont aucune signification...

- Tous les électeurs votent, et il n'y a pas de bulletin nul (en revanche, la méthode peut très bien prévoir un type de suffrage « blanc »). Cela sert uniquement à simplifier la présentation, puisqu'on peut toujours décider qu'il y a des types de suffrages « abstention » et « nul » ayant le même effet qu'un suffrage « blanc ».
- L'option vainqueur de l'élection dépend uniquement de la *proportion* d'électeurs ayant voté tel ou tel type de suffrage, pas de leur nombre absolu (par exemple, cela exclut que la méthode contienne une condition comme « seuls les types de suffrages ayant été votés par au moins 2 électeurs comptent »).
- On suppose enfin que les cas "limites" où **deux options sont rigoureusement à égalité**, ou où tout le monde a voté blanc, etc., ne se produisent jamais [5] : c'est en effet ce qui se passe en pratique quand le nombre d'électeurs est très grand, donc autant ne pas s'ennuyer à préciser ce qu'on fait en cas d'ex-æquo.

Le paradoxe de Condorcet

Le référendum : jusqu'ici tout va bien...

La majorité a toujours raison !

Une pensée du célèbre Chat de Philippe Geluck.

La problématique de la démocratie est souvent résumée par la maxime « la majorité a toujours raison », ce qui paraît très simple à première vue. Ainsi, si le choix doit se faire entre 2 options seulement (ce que nous appellerons un *référendum*), les cinq méthodes que nous avons données dans le « dilemme des Alliérins » de l'introduction se confondent en une seule et même méthode : c'est l'option que la majorité des électeurs préfèrent à l'autre qui gagne ! Nous appellerons cette méthode (spécifique aux référendums) la *méthode de la majorité*. On peut montrer que la méthode de la majorité est (dans un certain sens) la seule méthode démocratique "convenable" dans les situations de référendum : voir le **"théorème du référendum"** quelques paragraphes plus loin.

♪ Sitôt qu'on est plus de deux... ♪

Toutefois, cette belle simplicité s'écroule dès qu'il y a au moins 3 options ! La découverte de ce phénomène est due au mathématicien français du XVIIIe siècle **Nicolas de Condorcet**, qui fut le plus important précurseur de l'étude mathématique de la démocratie avec son *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* écrit en 1785 [6]. Dans cet ouvrage, Condorcet démontre que la méthode de la majorité peut aboutir à des incohérences, et que donc il n'est pas raisonnable de poser que la majorité ait toujours raison :

Nicolas de Condorcet

Portrait de Nicolas de Condorcet par Jean-Baptiste Greuze.

Paradoxe de Condorcet (1785) : *Il existe des situations où les préférences majoritaires des électeurs sont incohérentes, au sens où on peut trouver trois options X, Y et Z telles qu'une majorité d'électeurs préfèrent X à Y, une majorité d'électeurs préfèrent Y à Z, mais pourtant une majorité d'électeurs préfèrent Z à X.*

Démonstration (piste verte)

On peut imaginer que 40 % d'électeurs ont l'ordre de préférence « X puis Y puis Z », 35 % « Y puis Z puis X » et 25 % « Z puis X puis Y ». Alors 65 % des électeurs préfèrent X à Y, 75 % préfèrent Y à Z, mais pourtant 60 % préfèrent Z à X.

Dans les situations concrètes, il semblerait que le paradoxe de Condorcet n'arrive en fait que rarement [7]. Mais du point de vue théorique, il s'agit quand même d'une belle épine dans notre pied : s'il est possible de se retrouver face à des situations aussi contraires à l'intuition, y a-t-il une chance qu'on puisse définir une théorie mathématique cohérente de la volonté du peuple ?

Le problème de la manipulation : élections, piège à cons ?

Une méthode parfaite... ou presque

En fait, la raison qui explique que le problème de déterminer la volonté du peuple soit si délicat n'est elle-même pas vraiment mathématique, mais plutôt... biologique. Pour comprendre pourquoi, nous allons présenter une méthode électorale "idéale" qui résoudrait tous nos problèmes dans un monde parfait, puis voir pourquoi celle-ci ne convient en fait pas dans la réalité. Dans cette méthode (que nous noterons « F »), chaque électeur vote en donnant une note entre 0 et 20 à chaque option, « 20 » signifiant qu'il adhère totalement à cette option et « 0 » qu'il la rejette totalement. Pour le dilemme des Alliérins, par exemple, on pourrait imaginer que les électeurs donnent une note de 20 à leur propre ville et que pour les autres villes ils enlèvent 1 point tous les 5 kilomètres de distance, de sorte que les notes données par les électeurs seraient celles du tableau ci-dessous.

Les notes que les différents électeurs donneraient sincèrement aux différentes villes.

Provenance de l'électeur	Note de N	Note de S	Note de Y
N	20	5	1
S	5	20	8
Y	1	8	20

L'idée derrière cette méthode, c'est que la volonté du peuple doit être l'option qui optimise le bonheur total de la population, les notes servant à mesurer ce "bonheur". Du point de vue philosophique, cette définition de la volonté du peuple (appelée ***utilitarisme***) est certainement la plus raisonnable qu'on puisse donner ; et du point de vue mathématique, la méthode F permet de résoudre le paradoxe de Condorcet puisque l'ordre entre deux options est simplement l'ordre entre leurs notes moyennes et conduit donc à un classement sans ambiguïté [8]. Est-ce à dire que la méthode F est la méthode parfaite ? Hélas, non. Elle le serait... si Homo sapiens n'avait pas la capacité de mentir ! Voici en effet ce qui pourrait se passer dans la réalité, pour 40 % d'habitants de N, 35 % de Y et 25 % de S :

1. Tout d'abord, si les électeurs votaient sincèrement, les notes seraient celles du tableau ci-dessus. S gagnerait alors par une moyenne de 9,80 points, suivi de N (9,60 points) et enfin Y (9,40 points).
2. Mais comme les habitants de N sentent bien qu'ils risquent de finir seconds, ils peuvent décider de "mentir" en donnant à S une note de 0, bien que celle-ci ne corresponde pas à leur opinion véritable : dans ce cas en effet, la moyenne de S descend à 7,80 et c'est N qui gagne !
3. Anticipant cette manœuvre, les habitants de Y se disent alors que c'est eux qui risquent de finir seconds... Du coup, ils décident de mentir à leur tour en donnant à N une note de 0 : cela fait descendre la moyenne de N à 9,25 et c'est maintenant Y qui gagne !
4. Et, de stratagème en anticipation, on peut encore continuer longtemps comme cela...

Ce qui est sûr en tout cas, c'est que cette possibilité de voter un suffrage mensonger afin d'obtenir une vainqueuse qu'on préfère à celui que donnerait un suffrage sincère (ce que nous appellerons une *manipulation*) va complètement dévoyer la méthode en incitant les électeurs à donner des notes n'ayant que peu de rapport avec leur opinion véritable. Une anecdote raconte que le mathématicien français **Jean-Charles de Borda**, qui dans un ouvrage de 1770 avait défendu une méthode proche de la méthode F [9], répondit à Condorcet qui lui objectait la vulnérabilité de celle-ci à la manipulation : « ma méthode est faite pour des gens honnêtes ».

Si la maxime de Borda était pleine d'humanisme, force est malheureusement de constater qu'elle ne correspond guère à la réalité... Au contraire, il paraît plus réaliste de supposer que les électeurs n'hésiteront pas à utiliser dans leur intérêt personnel toute possibilité de manipulation permise par la méthode qui leur est proposée. Par conséquent, il paraît raisonnable d'exiger d'une méthode électorale, pour la qualifier de "bonne", qu'elle soit robuste à la manipulation, c'est-à-dire que les électeurs aient toujours intérêt à voter selon leur opinion véritable. Est-ce possible ?

Le théorème de Gibbard & Satterthwaite

Précisons la question que nous venons de poser. Il s'agit donc de savoir s'il existe une méthode démocratique pour laquelle les électeurs ont toujours intérêt à exprimer sincèrement leur opinion. Compte tenu que, comme nous venons de le voir, les méthodes de notation sont facilement vulnérables à la manipulation, il est clair que les seules questions qu'on pourra poser aux électeurs dans une méthode robuste à la manipulation devront concerner quelles options ils préfèrent à quelles autres, *sans indication de nuance*. Par conséquent, les types de suffrages possibles correspondront à un ordre de préférence entre les différentes options. Nous supposerons que les bulletins n'autorisent pas les ex-æquo, au motif qu'il y a toujours un chouïa qui fait basculer la préférence de l'électeur d'un côté ou de l'autre [10]. Ainsi, s'il y a 3 options X, Y et Z, il y aura 6 types de suffrages possibles qui sont « X puis Y puis Z », « X puis Z puis Y », « Y puis X puis Z », « Y puis Z puis X », « Z puis X puis Y » et « Z puis Y puis X ». Ce cadre étant précisé, la question de savoir s'il existe une méthode démocratique pour laquelle les électeurs ont toujours intérêt à voter sincèrement devient un problème mathématique parfaitement posé. Quelle est sa réponse ?

Mark Satterthwaite

Photographie contemporaine de l'économiste Mark Satterthwaite.

C'est en 1973 que deux chercheurs américains ont résolu indépendamment le problème... par la négative. Notez qu'aucun des deux n'appartenait au milieu des mathématiciens : Allan Gibbard était étiqueté comme philosophe et Mark Satterthwaite comme économiste... Mais c'est bien un théorème qu'ils ont démontré :

Théorème de Gibbard-Satterthwaite (1973) : *Supposons que les types de suffrages possibles correspondent à classer par ordre de préférence les différentes options. Alors, quelle que soit la méthode démocratique pour déterminer l'option vainqueur à partir des votes des électeurs, dès qu'il y a au moins 3 options, on peut trouver une situation dans laquelle certains électeurs ont intérêt à voter un suffrage qui ne reflète pas leur opinion véritable.*

Allan Gibbard

Photographie contemporaine du philosophe Allan Gibbard.

Démonstration (piste noire) [11]

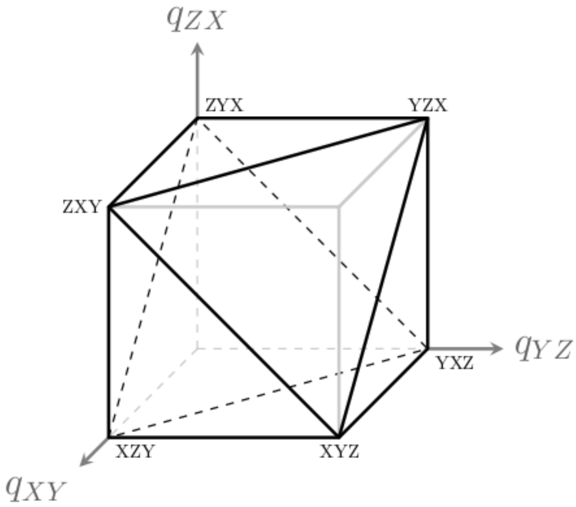
Supposons par l'absurde que nous ayons une méthode pour 3 options ou plus pour laquelle aucun électeur n'ait jamais intérêt à mentir. Nous considèrerons seulement le cas où il y a exactement trois options en lice X, Y et Z, vu qu'on peut toujours se ramener à ce cas en imaginant que ces trois options sont unanimement préférées à toutes les autres.

Imaginons que les proportions des différents types de suffrages sont telles que X soit déclarée vainqueuse, et considérons un électeur dont l'opinion véritable est « Y puis X puis Z ». Si cet électeur remplace son suffrage sincère par un suffrage mensonger « Y puis Z puis X », alors la vainqueuse après ce replacement ne doit pas pouvoir être Y, car sinon il serait *avantageux pour l'électeur de mentir, ce qui est interdit par notre hypothèse. Ainsi, remplacer un suffrage « Y puis X puis Z » par un suffrage « Y puis Z puis X » ne peut pas faire changer la vainqueuse de X pour Y. Cela ne peut pas non plus faire changer la vainqueuse de Y pour X (ce qui serait bizarre, du reste), car sinon en faisant marche arrière on ferait changer la vainqueuse de X pour Y en remplaçant un suffrage « Y puis Z puis X » par un suffrage « Y puis X puis Z », ce qui est interdit par le même argument.*

*Avec le même raisonnement, on peut également montrer qu'on ne peut pas non plus faire changer la vainqueuse de X pour Y ni de Y pour X en remplaçant un suffrage « Z puis Y puis X » par un suffrage « Y puis Z puis X », ni en remplaçant « X puis Z puis Y » par « X puis Y puis Z », ni en remplaçant « Z puis X puis Y » par « Z puis X puis Y » (ou l'inverse). Au final, on se rend compte qu'**aucune modification de suffrage ne peut faire changer la vainqueuse de X pour Y ni de Y pour X dès lors que cette modification ne change pas l'ordre relatif entre X et Y chez le suffrage concerné**.*

*On a évidemment la même chose en permutant les rôles de X, Y et Z : aucune modification de suffrage ne peut faire changer la vainqueuse entre X et Z dès lors qu'elle ne change pas l'ordre relatif entre X et Z chez le suffrage concerné, et aucune modification de suffrage ne peut faire changer la vainqueuse entre Y et Z dès lors qu'elle ne change pas l'ordre relatif entre Y et Z chez le suffrage concerné. Par conséquent, **la vainqueuse de l'élection dépend uniquement de la proportion de suffrages chez lesquels X est classée devant Y, de la proportion de suffrages où Y est classée devant Z et de la proportion de suffrages où Z est classée devant X**, puisque si aucune de ces trois proportions n'est modifiée, la vainqueuse ne peut changer ni entre X et Y, ni entre X et Z, ni entre Y et Z.*

Du coup, à toute valeur possible $(q_{XY}, q_{YZ}, q_{ZX}) \in [0,1]^3$ du triplet (« proportion de suffrages où X est classée devant Y », etc.), notre méthode associe une vainqueuse parmi {X,Y,Z}. On peut alors diviser l'espace tridimensionnel (ou plus précisément la partie de l'espace correpondant aux triplets possibles) en trois zones « X », « Y » et « Z », où la zone X est l'ensemble des points dont les coordonnées sont un triplet (q_{XY}, q_{YZ}, q_{ZX}) pour lequel X est déclarée vainqueuse, etc. : voir le dessin ci-dessous.



Représentation des préférences binaires entre X, Y et Z

L'ensemble des valeurs que peut prendre le triplet (q_{XY}, q_{YZ}, q_{ZX}) est un octaèdre irrégulier (en noir) obtenu en coupant deux coins opposés du cube $[0,1]^3$ (en gris). Les six sommets de cet octaèdre correpondent aux cas où tous les bulletins sont du même type : ainsi, le sommet « ZYX » correspond au cas où tous les bulletins sont du type « Z puis Y puis X », etc. Sur ce dessin, la coordonnée « arrière-avant » indique la proportion q_{XY}, la coordonnée « gauche-droite » indique q_{YZ} et la coordonnée « bas-haut » indique q_{ZX}.

Puisque faire changer la vainqueuse de X pour Z impose de modifier q_{ZX}, cela signifie que la frontière entre les zones X et Z doit correspondre (localement) à une valeur constante de q_{ZX}, autrement dit que cette frontière est horizontale. Or, quand q_{XY}, q_{YZ} et q_{ZX} sont toutes les trois égales, les trois options jouent des rôles complètement symétriques, de sorte qu'on se situe nécessairement à la frontière entre les trois zones à la fois ; par conséquent, la droite « $q_{XY} = q_{YZ} = q_{ZX}$ » est en particulier contenue dans la frontière entre les zones X et Z, et donc est horizontale. Mais cela n'est pas vrai vu que q_{ZX} varie le long de cette droite ! La seule manière d'éviter cette absurdité est de conclure que notre hypothèse de départ était fausse, autrement dit qu'aucune méthode électorale démocratique pour 3 options ou plus n'est robuste à la manipulation.

Notez que l'hypothèse qu'il y a au moins 3 options est cruciale, puisqu'en cas de référendum, la méthode de la majorité est robuste à la manipulation : c'est le "théorème du référendum", dont l'origine se perd dans la nuit des temps :

Théorème du référendum : *Supposons que nous sommes en présence d'un référendum (les options étant appelées « X » et « Y ») et qu'il y a deux types de suffrages possibles : « je préfère X » et « je préfère Y ». Alors :*

i. *La méthode de la majorité est robuste à la manipulation, c'est-à-dire que les électeurs ont toujours intérêt à exprimer leur opinion véritable ;*
ii. *Cette méthode de la majorité est la seule méthode démocratique qui soit robuste à la manipulation.*

Démonstration du premier point (piste verte)

Sous la méthode de la majorité, un électeur qui remplacerait son suffrage pour X par un suffrage pour Y ferait diminuer le score de X et augmenter celui de Y, de sorte qu'il pourrait éventuellement faire changer le résultat d'une victoire de X vers une victoire de Y, mais pas l'inverse. Par conséquent, si l'opinion véritable de cet électeur est en faveur de X, il n'a jamais intérêt à faire ce remplacement. De même, un électeur dont l'opinion véritable est en faveur de Y n'a jamais intérêt à voter&pour;X, ce qui prouve le point i.

Démonstration du second point (piste bleue)

Pour le point ii, supposons que nous avons une méthode démocratique robuste à la manipulation ; notre objectif est de montrer que cette méthode est nécessairement la méthode de la majorité. La robustesse à la manipulation implique que quand un électeur remplace un suffrage pour Y par un suffrage pour X, cela peut éventuellement faire gagner Y à la place de X, mais pas l'inverse (car sinon un électeur dont l'opinion véritable est en faveur de X aurait intérêt à voter un suffrage mensonger pour Y). En conséquence, si on passe progressivement de 0 % de suffrages pour Y (et 100 % pour X) à 100 % de suffrages pour Y (et 0 % pour X), il se peut qu'à un moment la vainqueuse passe de X à Y, mais pas l'inverse. Du coup, au cours de ce passage progessif de 0 % de suffrages pour Y à 100 %, ce qui se produit est forcément qu'il existe un seuil en deçà duquel c'est X qui gagne et au-delà duquel c'est Y qui gagne.

Notons S le pourcentage correspondant à ce seuil, c'est-à-dire que si Y reçoit plus de S % des suffrages, c'est Y qui gagne, et que sinon c'est X qui gagne. Comme les options X et Y jouent le même rôle, on a aussi que X gagne s'il reçoit plus de S % des suffrages et que Y gagne sinon. Le fait que Y reçoive plus de S % des suffrages doit donc être équivalent au fait que X reçoive moins de S % des suffrages ; par conséquent les deux options doivent se retrouver à exactement S % des suffrages au même moment. À ce moment-là, le total des suffrages sera alors de (2×S) %, et comme ce total des suffrages doit de toutes façons être de 100 %, c'est que nécessairement S est égal à 50 : on en conclut que notre méthode est bien la méthode de la majorité, ce qu'on voulait.

Le théorème de Gibbard-Satterthwaite s'inscrit dans une lignée de grands résultats négatifs établis au cours de la seconde moitié du XXe siècle. La découverte de ces résultats fut une petite révolution théorique : ainsi, **le premier des résultats de cette lignée**, établi par l'économiste américain Kenneth Arrow en 1952, lui valut la quatrième édition du **"prix Nobel d'économie"** vingt ans plus tard !

Kenneth Arrow

Kenneth Arrow, “prix Nobel d’économie” 1972 (photographié en 2004).

Tout cela est tout de même assez déprimant : le seul intérêt de l’étude mathématique de la démocratie serait-il donc de gagner 240 000 couronnes [12] ? La théorie de la volonté du peuple serait-elle nécessairement absurde ? Restons optimistes : maintenant que nous avons touché le fond, nous ne pouvons plus que remonter... Et de fait, il y a aussi des résultats *positifs* dans la théorie de la volonté populaire ! Nous allons voir un tel résultat pas plus tard que tout de suite.

Le scrutin stochocratique, remède à tous nos maux ?

Attaquons ce paragraphe par une annonce fracassante : contrairement à ce que je vous ai laissé entendre, il existe bel et bien une méthode démocratique pour laquelle aucun électeur n’a jamais intérêt à voter de façon mensongère ! « Comment cela est-il possible sans contrevenir au théorème de Gibbard-Satterthwaite ? » me demanderez-vous. Eh bien, le fait est que j’ai passé sous silence une hypothèse implicite dans ce théorème... À savoir, j’ai supposé que l’option vainqueur de l’élection devait être déterminée *uniquement* par les votes des électeurs. Cette hypothèse paraît nécessaire à première vue, puisqu’on se dit que faire intervenir un paramètre extérieur romprait l’égalité entre les options ou entre les électeurs... Mais il y a tout de même un paramètre extérieur qui ne rompe pas cette égalité : le *hasard* ! En d’autres termes, on peut imaginer que vote des électeurs ne désigne pas *directement* l’option vainqueur de l’élection, mais que celle-ci soit désignée in fine par un *tirage au sort* — tirage dont, bien sûr, les règles dépendront des votes exprimés par les électeurs (ainsi, si par exemple tous les électeurs ont classé les deux mêmes options devant toutes les autres, il paraît raisonnable que dans ce cas le tirage ne permette la victoire que d’une de ces deux options-là).

On peut ainsi imaginer une méthode où chaque électeur écrit sur son bulletin son option préférée, puis où on mélange tous les bulletins avant d’en tirer un au hasard et de décider que celui-ci désigne l’option vainqueur de l’élection : nous appellerons cette méthode le *scrutin stochocratique*. Bien que cette méthode fasse appel au hasard, elle n’est pas complètement arbitraire non plus dans la mesure où, par exemple, une option n’a aucune chance de gagner si tous les électeurs lui préfèrent une même autre option : c’est ce qu’on appelle le *critère d’unanimité*. En outre, cette méthode est robuste à la manipulation :

Théorème : *Dans un scrutin stochocratique, tous les électeurs ont toujours intérêt à exprimer leur opinion véritable.*

Démonstration (piste verte)

Quand un électeur réfléchit au suffrage qu’il va voter, il doit prendre en compte deux éventualités, dont les probabilités ne dépendent pas du choix qu’il va faire :

- *Soit le bulletin tiré au sort est celui d’un autre électeur. Dans ce cas le suffrage de notre électeur n’est pas pris en compte, et il n’a donc aucun intérêt à mentir.*
- *Soit le bulletin tiré au sort est le sien, et dans ce cas notre électeur a tout intérêt à exprimer son opinion véritable.*

Dans tous les cas, il vaut donc mieux pour l’électeur exprimer son opinion véritable.

Le scrutin stochocratique est donc une méthode robuste à la manipulation. Cela dit, c’est tout de même une méthode assez “brutale”, car faisant courir le risque, bien que faible, que gagne une option qui déplaît à la grande majorité des électeurs ; en particulier, quand il n’y a que 2 options, cette méthode ne redonne pas la méthode de la majorité ! On aimerait donc bien disposer d’une autre méthode du même type qui soit plus “raisonnable” de ce point de vue... Mais il se trouve hélas que le scrutin stochocratique est la seule méthode qui satisfasse les critères que nous venons de voir :

Théorème “de la dictature aléatoire” (Gibbard, 1977) : *Supposons que les types de suffrages possibles correspondent à classer par ordre de préférence les différentes options. Alors, dès qu’il y a au moins 3 options, la seule méthode démocratique avec tirage au sort éventuel qui satisfasse le critère d’unanimité (càd. pour laquelle une option unanimement classée derrière une même autre ne gagne jamais) et pour laquelle aucun électeur n’ait jamais intérêt à mentir est le scrutin stochocratique.*

Démonstration (piste noire) [13]

Supposons que nous disposons d’une méthode robuste à la manipulation et satisfaisant le critère d’unanimité ; nous allons montrer que cette méthode est nécessairement le scrutin stochocratique. Pour simplifier la présentation, nous nous limitons au cas où seulement 3 options sont en lice, notées X, Y et Z ; sachant que le raisonnement que nous allons suivre s’adapte directement au cas d’un nombre arbitraire d’options [14]. Nous abrègerons les types de suffrages « X puis Y puis Z », « X puis Z puis Y », etc. en « XYZ », resp. « XZY », etc. Nous noterons p_X, p_Y et p_Z les probabilités que X, Y et Z gagnent à l’issue du tirage au sort, ces probabilités dépendant déterministiquement de la proportion de suffrages des différents types.

Commençons par observer que, quand un électeur remplace un suffrage XYZ par un suffrage XZY, cela ne peut pas faire varier p_X (et de même, remplacer un suffrage YXZ par YZX ne pourra pas faire varier p_Y, ni remplacer ZXY par ZYX faire varier p_Y). En effet, imaginons un électeur aux yeux desquels Y et Z ont pratiquement la même valeur, mais qui estime que X leur est très supérieure. Pour cet électeur, la seule chose qui compte vraiment est de maximiser p_X ; par conséquent, s’il arrivait à faire augmenter p_X en remplaçant son suffrage XYZ par XZY, cela servirait son intérêt même si en réalité il préfère d’un chouïa Y à Z ; or cela est exclu par l’hypothèse de robustesse à la manipulation. C’est donc que remplacer un suffrage XYZ par XZY ne peut pas faire augmenter p_X ; et cela ne peut pas le faire diminuer non plus puisque sinon c’est remplacer un suffrage XZY par XYZ qui ferait augmenter p_X, ce qu’on exclut de la même façon.

*Nous venons ainsi d’établir que permuter les deux dernières options d’un suffrage ne peut pas faire varier la probabilité de victoire de l’option classée première sur ce suffrage. On montrerait de la même façon que permuter les deux **premières** options d’un suffrage ne peut pas faire varier la probabilité de victoire de l’option classée **dernière** sur ce suffrage : par exemple, remplacer XYZ par YXZ ne peut pas faire varier p_Z. [15]*

*Maintenant, imaginons la situation suivante. Dans un premier temps, un électeur votant XYZ change d’avis et décide de voter XZY : d’après ce que nous venons de voir, cela a pour seul effet éventuel d’augmenter p_Z d’une probabilité ε (qui pourrait être en fait négative sans perturber le raisonnement) et de diminuer p_Y de la même quantité. Puis dans un second temps, une autre électrice change à son tour d’avis en remplaçant, pour sa part, son suffrage ZXY par ZYX : d’après ce que nous venons de voir, cela ne fait pas varier p_Z. L’effet combiné de ces deux changements sur p_Z est donc une augmentation de ε. Regardons maintenant ce qui se passe quand on inverse l’ordre des changements d’avis, c’est-à-dire que la seconde électrice change d’avis en première : alors dans un premier temps p_Z ne varie pas, et dans un second temps le remplacement du suffrage XYZ par XZY fait augmenter p_Z de η (et diminuer p_Y de la même quantité). A priori ε et η pourraient être différents, vu que la répartition des suffrages au moment du changement d’avis de l’électeur de XYZ n’est pas la même dans la nouvelle situation que dans l’ancienne. Mais l’effet combiné des changements d’avis doit être le même dans les deux cas, de sorte qu’en fait ε et η sont nécessairement égaux. On a ainsi montré la chose suivante : **l’effet du remplacement d’un suffrage XYZ par un suffrage XZY est le même quand, parmi les autres suffrages, on remplace un ZXY par un ZYX**.*

*On peut montrer de la même façon que plus généralement, l’effet du remplacement d’un suffrage XYZ par un suffrage XZY est le même quand on permute sur un des autres suffrages deux options **adjacentes** qui ne sont pas **collectivement** Y et Z (c’est-à-dire qu’une des deux options permutées peut être Y ou Z, mais pas les deux). En répétant plusieurs fois cette opération, on peut considérablement changer la répartition des autres suffrages : la seule contrainte sur le changement de cette répartition est que le nombre des autres suffrages plaçant Y devant Z (ici, quand je dis « devant », cela signifie « devant “tout court” », pas « juste devant ») doit rester le même. Finalement, **l’effet du remplacement d’un suffrage XYZ par un suffrage XZY ne dépend que de la proportion de suffrages chez lesquels Y est classée devant Z** ; de même, l’effet du remplacement d’un suffrage YXZ par YZX ne dépend que de la proportion de suffrages classant X devant Z, etc.*

*On peut faire le même raisonnement en permutant non pas les deux dernières options, mais les deux premières options : on obtient alors que **l’effet du remplacement d’un suffrage XYZ par un suffrage YXZ ne dépend que de la proportion de suffrages chez lesquels X est classée devant Y** ; de même, l’effet du remplacement d’un suffrage XZY par ZXY ne dépend que de* *la proportion de suffrages classant X devant Z, etc.*

Ayant exploré les conséquences du critère de robustesse à la manipulation, voyons maintenant une conséquence du critère d'unanimité : ***si tous les suffrages classent X en tête, alors X doit être élu avec une probabilité de 100 %****. En effet, si tous les suffrages placent X en tête, cela signifie en particulier qu'ils placent tous X devant Y, et donc Y ne doit pas pouvoir être élu, ni Z d'après le même raisonnement, de sorte que c'est nécessairement X qui doit gagner.*

Justement, regardons ce que donnent les conséquences du critère de robustesse quand tous les suffrages placent X en tête, c'est-à-dire que les seuls suffrages votés ont été XYZ et XZY. Dans ce cas, d'après le paragraphe précédent, remplacer un suffrage XYZ par XZY n'a aucun effet puisque c'est toujours X qui gagne à 100 % ! Or, de manière générale, nous savons que l'effet du remplacement d'un suffrage XYZ par XZY ne dépend que de la proportion de suffrages où Y est classée devant Z. Comme n'importe quelle proportion de suffrages où Y est classée devant Z peut être atteinte par une situation où tous les suffrages placent X en tête, on en déduit le résultat suivant : ***remplacer un suffrage XYZ par un suffrage XZY n'a jamais aucun effet !*** *Évidemment, X ne joue aucun rôle particulier dans ce raisonnement, donc on en déduit que permuter les deux dernières options sur un suffrage n'a jamais aucun effet ; autrement dit :* p_X, p_Y ***et*** p_Z ***ne dépendent que des options classées en tête par les suffrages****.*

Vu que seules les options classées en tête comptent, nous allons dorénavant parler de « suffrages pro-X » pour dire « suffrages où X est classée en tête », etc. La question que nous nous posons maintenant est la suivante : quel est l'effet du remplacement d'un suffrage pro-X par un suffrage pro-Y ? On peut imaginer que cela correspond en fait au remplacement d'un suffrage XYZ par un suffrage YXZ, de sorte que d'après ce que nous avons dit précédemment, l'effet que cela a sur p_X, p_Y *et* p_Z *ne dépend que de la proportion de suffrages où X est classée devant Y. Mais nous savons* ***aussi****, d'après le paragraphe précédent, que cet effet ne dépend que des proportions des suffrages pro-X, pro-Y et pro-Z ! Or, il est possible de passer de n'importe quelle proportion de suffrages à n'importe quelle autre en ne procédant qu'aux deux opérations suivantes :*

- *Soit modifier les préférences de chaque suffrage sans jamais changer le classement relatif de X et Y ;*
- *Soit modifier les préférences de chaque suffrage sans jamais changer l'option classée en tête.*

Au final, puisqu'aucune de ces deux opérations ne modifie l'effet du remplacement d'un suffrage pro-X par un suffrage pro-Y, on en déduit que ***l'effet du remplacement d'un suffrage pro-X par un suffrage pro-Y ne dépend absolument pas de la répartition des suffrages****.*

Nous y sommes presque ! Grâce à ce que nous venons d'établir, on a que quand on passe progressivement de 100 % de suffrages pro-X (où, d'après le critère d'unanimité, on a $p_X = 100$ *%,* $p_Y = 0$ *% et* $p_Z = 0$ *%) à 100 % de suffrages pro-Y (où on a de même* $p_X = 0$ *%,* $p_Y = 100$ *% et* $p_Z = 0$ *%), chaque remplacement d'un suffrage pro-X par un suffrage pro-Y a le même effet ; donc, s'il y a* q_X *% de suffrages pro-X, (100−*q_X*) % de suffrages pro-Y et 0 % de suffrages pro-Z, on a* $p_X = q_X$ *%,* $p_Y = (100-q_X)$ *% et* $p_Z = 0$ *%. Si dans un second temps nous modifions la proportion de suffrages pro-Y et pro-Z, on peut imaginer que cela se fait en remplaçant des suffrages YZX par des suffrages ZYX, ce qui ne change pas* p_X *comme nous l'avons établi plus haut : ainsi, après ce changement, la probabilité que X soit vainqueur est toujours exactement égale à la proportion* q_X *de suffrages pro-X. Il en va évidemment de même pour Y et pour Z, donc au final* ***la probabilité pour une option d'être vainqueur est exactement égale au nombre de suffrages classant cette option en tête****, ce qui signifie bien que notre méthode est le scrutin stochocratique.*

Les recherches continuent

Résumons. La problématique de déterminer la volonté du peuple à partir des préférences individuelles n'a pas de solution évidente a priori, en particulier à cause du fait que les électeurs ont la possibilité de manipuler le scrutin en exprimant une opinion mensongère afin de mieux faire triompher leur camp. Quand il n'y a que 2 options en lice, la méthode de la majorité surmonte cette difficulté et s'impose comme la méthode la plus convenable ; mais dès qu'il y a au moins 3 options, le paradoxe de Condorcet sur l'incohérence des préférences majoritaires empêche l'existence d'une telle méthode "parfaite". La seule façon d'assurer que tous les électeurs votent sincèrement est le scrutin stochocratique, mais celui-ci peut parfois désigner une vainqueuse manifestement non satisfaisante !

Alors, que faire ? Faut-il préconiser le scrutin stochocratique malgré tout ? Cela ne semblerait vraiment pas raisonnable [16]... Faut-il abandonner la démocratie en raisons de ses paradoxes ? Cela serait tout de même excessif...

Les chercheurs n'ont pas baissé les bras. Ils se sont dits : « il n'existe pas de méthode électorale parfaite ? Eh bien, nous allons chercher une méthode "aussi parfaite que possible", qui ne satisfasse peut-être pas tous les critères souhaitables mais qui en satisfait malgré tout un grand nombre ! Il n'existe pas de bonne règle générale pour déterminer le vainqueur dans toutes les situations ? Eh bien, nous allons chercher une règle "aussi générale que possible" qui donne un vainqueur incontestable dans la plupart des cas réalistes ! ». C'est sur ces passionnantes problématiques que nous nous pencherons dans la suite de cette série d'articles.

Références

Le lecteur désireux d'aller plus loin pourra consulter les références suivantes. Attention, celles-ci sont de difficulté « piste noire », voire hors piste !

- *Social Choice and the Mathematics of Manipulation*, par Alan D. Taylor (2005) ; Cambridge University Press.
- *A geometric proof of Gibbard's random dictatorship theorem*, par John Duggan (1996) ; *Ecomonic Theory* no. 7, pp. 365–369 ; Springer-Verlag.
- **wiki.electorama.com** : Un wiki créé par les membres du groupe de discussion « Election-Methods ».

P.S. :

Je tiens à remercier les nombreuses personnes qui m'ont aidé dans la rédaction de cet article. J'ai grandement apprécié la cordialité des membres du groupe de discussion « Election-Methods » (trop nombreux pour être listés ici) qui m'ont aimablement indiqué des références pertinentes sur le sujet ou m'ont donné leur écho de spécialistes sur mon texte. Les commentaires des relecteurs désignés par Images des Mathématiques, notamment Frédéric Chardard et Jean Lécureux, ont conduit à des améliorations substantielles de l'article. Je remercie également mes parents Benoît et Odile qui m'ont servi de cobayes pour vérifier si le texte était compréhensible par les non-mathématiciens, et mon infographiste de sœur Laurence pour ses conseils avisés qui m'ont permis de mettre au point la carte de l'Allier et le logo !

Notes

[1] Et le domaine est toujours bien vivant : ainsi, sur le groupe de discussion électronique « Election-Methods », les chercheurs s'échangent plus de 200 messages par mois !

[2] Un « type de suffrage » n'est pas la même chose qu'une « option » : les options sont les possibilités parmi lesquelles il faut choisir, mais l'information que contient un bulletin de vote sera en général plus complexe qu'une seule option... Par exemple, dans le cas de la "méthode B synthétisée en un seul scrutin", s'il y a 4 options, chaque électeur indique sur son bulletin, d'une part qui il choisit au premier tour, soit 5 possibilités (car il a aussi le droit de voter blanc), puis, pour chacun des 6 seconds tours possibles, qui il choisirait au second tour, soit 3 possibilités à chaque fois : il y a donc 5×3×3×3×3×3×3 = 3 645 types de suffrages possibles.

[3] Dans certains scrutins, on peut cependant souhaiter que certains électeurs aient plus de poids que d'autres : par exemple, dans un concours de chant comme *Nouvelle Star*, on pourrait vouloir faire voter en même temps le public et le jury d'experts, mais en donnant plus de poids à un juré qu'à un spectateur anonyme... En fait, ce genre de situations peut se ramener à celle où tous les électeurs ont le même poids, en disant qu'un électeur "privilégié" (en l'occurrence, un juré) compte pour plusieurs personnes à lui tout seul.

[4] Si nous étions en piste noire, nous dirions un peu pompeusement : « la méthode électorale est invariante par l'action de n'importe quelle permutation sur les électeurs ».

[5] En piste noire, nous dirions que « nous demandons seulement que la vainqueuse de notre élection soit définie sur un ensemble de comesure nulle dans l'espace des dépouillements » — l'espace des dépouillements étant, s'il y a k types de suffrages possibles, le $(k-1)$simplexe formé par les k-uplets positifs de somme 1 (chaque entrée d'un tel k-uplet correspondant à la proportion de suffrages d'un certain type), muni de la mesure de Lebesgue.

[6] On s'est aperçu récemment que plusieurs des découvertes de Condorcet avaient en fait déjà été faites au... XIIIe siècle par le savant majorquin **Raymond Lulle**, dont le manuscrit *Ars eleccionis*, écrit en 1299, a été redécouvert en 2001 après sept siècles d'oubli !

[7] Par exemple, dans le cas du dilemme des Alliérins de l'introduction, il n'y a jamais de paradoxe de Condorcet, quelle que soit la proportion d'habitants des différentes villes : nous verrons pourquoi dans le deuxième article de cette série.

[8] L'idée nouvelle apportée par l'utilitarisme et qui permet de résoudre le paradoxe de Condorcet, c'est qu'une option préférée par la majorité peut ne pas correspondre à la volonté du peuple si la majorité n'est convaincue que de justesse par cette option alors que la minorité la trouve franchement mauvaise

[9] En fait, la méthode que Borda proposait était exactement la méthode E.

[10] Au pire, si l'électeur n'arrive vraiment pas à se décider, il peut toujours tirer au sort entre les deux manières de départager l'ex-æquo... Comme le nombre d'électeurs est très grand, cela revient en pratique au même que s'il votait un demi suffrage de chaque type.

[11] Cette démonstration est sensiblement différente des preuves "classiques" qu'on peut trouver de ce résultat, parce qu'ici nous supposons que la proportion de suffrages de tel ou tel type est un paramètre continu, alors que d'habitude le théorème de Gibbard-Satterthwaite est énoncé pour un nombre fini d'électeurs ; en outre, nos hypothèses de symétrie entre électeurs et entre options sont nettement plus fortes ici que dans le "vrai" théorème de Gibbard-Satterthwaite.

[12] Soit environ 180 000 euros d'aujourd'hui.

[13] Cette démonstration est inspirée d'un article de John Duggan (cf. **références**).

[14] Notez que par contre, contrairement à la démonstration du théorème de Gibbard-Satterthwaite, on ne peut pas arguer que démontrer le résultat pour 3 options suffit à montrer le cas général, attendu qu'il ne s'agit pas d'arriver à une contradiction mais de prouver qu'il n'y a qu'une seule méthode qui marche.

[15] Dans le cas d'un nombre d'options quelconque, ce qu'on dit est que permuter deux options *adjacentes* ne peut pas faire varier la probabilité de victoire des autres options.

[16] Il est à noter cependant que dans la démocratie athénienne, mère de toutes les démocraties, la plupart des magistrats étaient désignés par tirage au sort (comme le sont encore aujourd'hui, à bien moindre échelle, les jurés de cour d'assises), et que les élections étaient même considérées en Grèce antique comme un procédé *non* démocratique, au motif qu'elles favorisaient les citoyens les plus riches ou les plus célèbres ! Aristote écrivait ainsi dans sa *Politique* (livre VI, chap. VII) : « la voie du sort pour la désignation des magistrats est une institution démocratique ; le principe de l'élection, au contraire, est oligarchique ».

Affiliation de l'auteur

Rémi Peyre : Institut Élie Cartan (Université de Lorraine) / CNRS

Pour citer cet article : **Rémi Peyre**, **« La démocratie, objet d'étude mathématique »** — *Images des Mathématiques*, CNRS, 2012. En ligne, URL : http://images.math.cnrs.fr/La-democratie-objet-d-etude.html

Les mathématiques de la démocratie, II

Et le vainqueur du second tour est...

Le critère de Condorcet

Le 10 mai 2012, par **Rémi Peyre**

Maître de conférences à l'École des Mines de Nancy. (**page web**)

... François Bayrou !

Certes, le candidat centriste n'est arrivé qu'en cinquième position au premier tour, mais s'il avait accédé au second, il y aurait vaincu (à en croire les sondages) n'importe lequel de ses concurrents... Au XVIII^e siècle, le Marquis de Condorcet a théorisé que, lorsqu'un candidat dans la situation de Bayrou existe, une méthode électorale bien conçue devrait nécessairement le désigner comme vainqueur : « s'il existe un candidat qui, lorsqu'on le confronte à n'importe quel autre candidat, est préféré à cet autre candidat par une majorité d'électeurs, alors ce candidat est celui d'entre tous que le peuple préfère ».

Dans ce texte, nous commencerons par expliquer quelles sont les justifications philosophiques et mathématiques de ce critère, puis nous regarderons dans quelles circonstances il y a ou pas un « vainqueur de Condorcet », avant de présenter une méthode qui généralise le critère de Condorcet lorsqu'aucun tel vainqueur n'existe. (Ce texte s'inscrit dans la continuité d'un article précédent du même auteur).

Préambule : Retour sur l'élection présidentielle française

Le 6 mai dernier, François Hollande a été élu Président de la République Française. Pour nos lecteurs étrangers, rappelons d'abord comment fonctionne une élection présidentielle en France. Un nombre restreint de candidats (10 en l'occurrence) de bords politiques variés ayant été préalablement sélectionnés par les élus locaux, on procède entre eux à un *scrutin uninominal à deux tours*. Dans un tel scrutin, on organise d'abord un premier tour de vote dans lequel chaque électeur doit choisir un seul des candidats (ou s'abstenir). Si un des candidats reçoit la majorité absolue des suffrages exprimés (ce qui est rare en pratique), alors ce candidat est déclaré vainqueur. Sinon, on procède à un second tour de vote dans lequel les deux candidats ayant reçu le plus de suffrages au premier tour s'affrontent en face-à-face ; le candidat recevant le plus de voix au second tour est alors déclaré vainqueur.

Rappelons également quels ont été les résultats de l'élection de 2012. Pour simplifier la présentation, nous ne retiendrons ici que les cinq candidats principaux (que nous abrègerons parfois par leur initiale), à savoir François BAYROU (B), François HOLLANDE (H), Marine LE PEN (L), Jean-Luc MÉLENCHON (M) et Nicolas SARKOZY (S). Les résultats du premier tour ont été les suivants [1] :

1. HOLLANDE (31 %)
2. SARKOZY (28 %)
3. LE PEN (19 %)
4. MÉLENCHON (13 %)
5. BAYROU (9 %)

Aucun des candidats n'ayant dépassé 50 % des voix, il y a donc eu un second tour opposant Hollande à Sarkozy, dont les résultats ont été les suivants :

1. HOLLANDE (52 %)
2. SARKOZY (48 %)

Au final, c'est donc Hollande qui a été élu président.

Soyons maintenant un peu plus curieux, et demandons-nous quels auraient pu être les résultats des autres faces-à-faces imaginables. Outre le second tour H – S qui a réellement eu lieu, il y aurait eu 9 autres possibilités : B – H, B – L, B – M, B – S, H – L, H – M, L – M, L – S et M – S. Évidemment nous ne saurons jamais avec certitude ce qu'auraient donné ces seconds tours fictifs ; toutefois, à en croire les instituts de sondage [2], les résultats auraient été à peu près les suivants [3] :

François Hollande

Hollande, le candidat élu par le peuple.

- BAYROU bat HOLLANDE 51 – 49 ;
- BAYROU bat LE PEN 74 – 26 ;
- BAYROU bat MÉLENCHON 75 – 25 ;
- BAYROU bat SARKOZY 55 – 45 ;
- HOLLANDE bat LE PEN 65 – 35 ;
- HOLLANDE bat MÉLENCHON 80 – 20 ;
- LE PEN bat MÉLENCHON 53 – 47 ;
- LE PEN est battue par SARKOZY 32 – 68 ;

- MÉLENCHON est battu par SARKOZY 36 – 64.

Nous synthétisons tout cela sous la forme du « diagramme des préférences binaires » ci-dessous, où :

- Chaque candidat est représenté par une position distincte ;
- Entre chaque paire de candidats, il y a une flèche qui pointe vers celui des deux qui serait vainqueur en face-à-face (on néglige la possibilité d'un ex æquo) ;
- Sur chaque flèche, on indique le score (en pourcentage) par lequel le vainqueur du face-à-face concerné gagne.

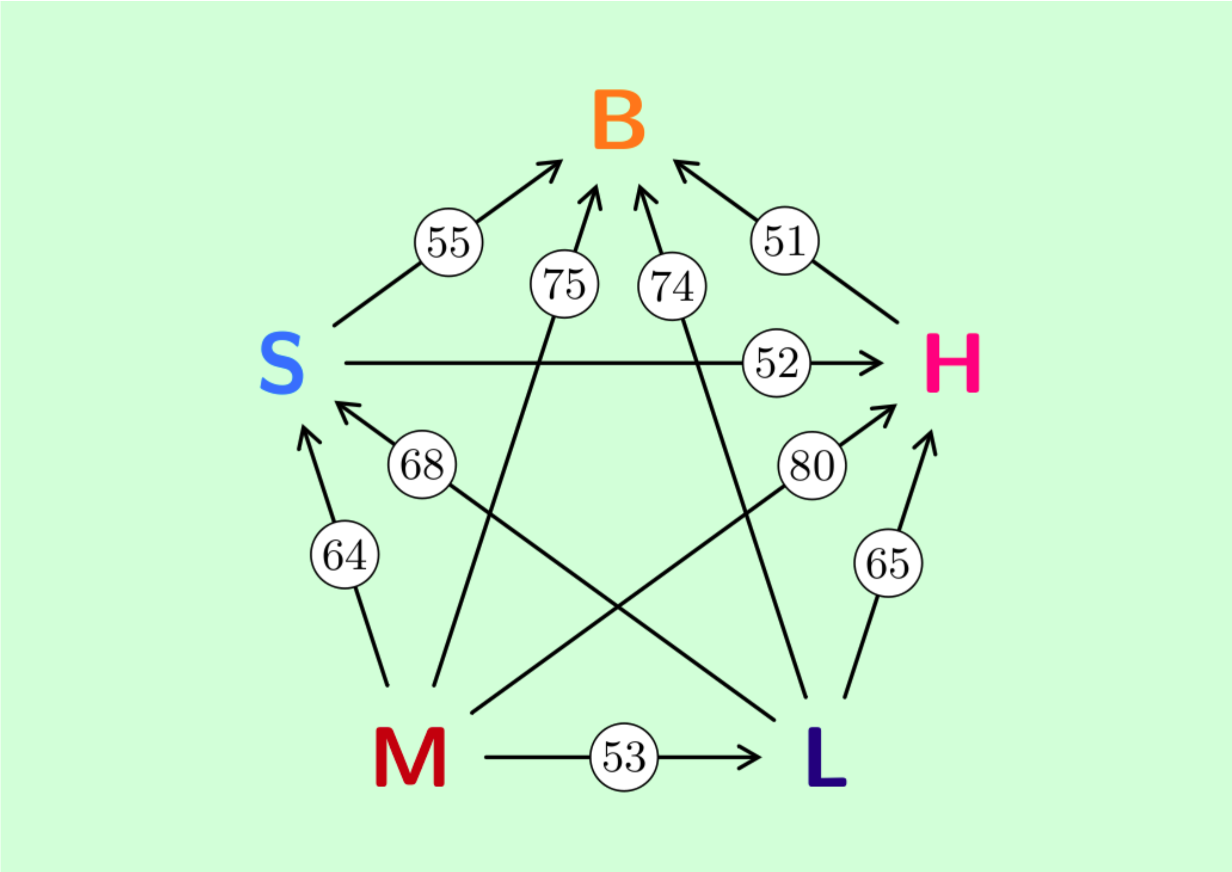


Diagramme des préférences binaires pour l'élection présidentielle

Ce qu'auraient pu donner les 10 seconds tours imaginables entre les cinq candidats principaux à l'élection présidentielle française de 2012 (les flèches pointent vers les vainqueurs).

François Bayrou

Bayrou, le candidat voulu par le peuple ?

Une constatation frappe : François Bayrou aurait vaincu n'importe lequel de ses adversaires au second tour ! Comme, dans la situation historique, il n'y a pas été qualifié, il n'a évidemment pas été élu, mais plusieurs éditorialistes [voir par exemple **ici** ou **là**] ont souligné qu'il y avait là une forme de paradoxe... Ne serait-il pas logique, en effet, qu'un candidat globalement préféré à n'importe quel autre soit considéré comme le préféré de tous [4] ?

Philosophie du critère de Condorcet

Énoncé du critère

C'était en tout cas l'opinion de **Nicolas de Condorcet**, savant français du XVIII^{e} siècle qui fut un des pionniers de l'étude mathématique de la démocratie. Il introduisit les définitions suivantes (que nous adaptons ici en langage moderne) :

Définitions (Vainqueur de Condorcet ; Critère de Condorcet)

> *Si, parmi les candidats à une élection, il en existe un qui, face à n'importe quel autre, lui est préféré par une majorité d'électeurs, alors ce candidat est appelé le vainqueur de Condorcet.*
>
> *On dit qu'une méthode électorale satisfait le critère de Concorcet quand, lorsqu'il y a un vainqueur de Condorcet, c'est toujours lui que cette méthode déclare vainqueur (sous réserve que les électeurs aient voté selon leurs préférences véritables).*

Nicolas de Condorcet

Portrait de Nicolas de Condorcet par Jean-Baptiste Greuze.

Avant toute chose, une vérification s'impose : nous parlons *du* vainqueur de Condorcet comme s'il était unique, mais est-ce bien le cas ? La réponse est « oui » :

Théorème (Unicité du vainqueur de Condorcet)

> *Il ne peut pas y avoir plus d'un vainqueur de Condorcet à une élection.*

Démonstration

> *Imaginons une situation dans laquelle il y aurait deux vainqueurs de Condorcet (au moins), appelés X et Y. Puisque X est un vainqueur de Condorcet, il vainc n'importe lequel de ses concurrents en face-à-face, donc en particulier Y. Mais de même, puisque Y est un vainqueur de Condorcet, il doit vaincre X en face-à-face... Or le face-à-face entre X et Y ne peut évidemment avoir qu'un seul vainqueur : c'est donc qu'il est en fait impossible qu'il y ait plus d'un vainqueur de Condorcet.*

La thèse de Condorcet était que, pour qu'une méthode électorale soit vraiment juste, il faut qu'elle satisfasse le critère éponyme. Citons le grand homme :

> Il peut réfulter de la manière de voter employée dans les élections ordinaires une décifion réellement contraire à la pluralité. Ainfi l'on devroit fubftituer à cette forme celle dans laquelle chaque Votant, exprimant l'ordre fuivant lequel il place les Candidats, prononceroit à la fois fur la préférence refpective qu'il leur accorde. Lorfqu'on voudra choifir le candidat le plus digne, il fuffira que le fyftème n'implique point contradiction pour le candidat qui mérite la préférence fur tous. [5]

Condorcet se serait donc opposé à notre méthode uninominale à deux tours, puisque l'exemple de Bayrou montre que celle-ci ne satisfait *pas* son critère ! Mais avant d'aller plus loin, il convient d'expliquer quelles sont les justifications théoriques à la thèse de Condorcet.

Justification mathématique du critère de Condorcet [6]

Si vous avez lu **mon article précédent**, vous vous souvenez sans doute qu'un grand problème de la théorie mathématique des élections est le suivant : les électeurs ont parfois intérêt à "voter stratégique" en exprimant une opinion mensongère qui amènerait le résultat du scrutin à être *meilleur* pour eux que celui qu'aurait donné un suffrage sincère ! Dans le cas du scrutin uninominal à deux tours dont nous parlions dans le préambule, un exemple typique est ce qu'on appelle le « vote utile ». Imaginons ainsi qu'à la veille du premier tour les sondages donnent Sarkozy loin en tête, suivi de Le Pen et Hollande au coude-à-coude, et nettement plus loin Mélenchon. Considérons un électeur fictif qui est grand supporteur de Mélenchon, est mitigé sur Hollande, n'aime pas Sarkozy et *déteste* Le Pen. Cet électeur se dit : « si je vote Mélenchon, cela ne suffira pas à le qualifier pour le second tour de toute façon ; par contre je cours le risque de voir un second tour entre les deux candidats que j'aime le moins. En revanche, en votant Hollande, je diminue les chances d'un second tour Sarkozy – Le Pen et je favorise l'élimination de ma candidate détestée ! J'ai donc intérêt à voter Hollande ».

Or, quand on prend en compte ce phénomène de vote stratégique, le vainqueur de Condorcet (s'il y en a un) peut s'arranger pour être élu à coup sûr ! (en tout cas dans la méthode uninominale à deux tours [7]). En effet, imaginons qu'il y a un vainqueur de Condorcet appelé X, et supposons que les sondages prédisent la victoire d'un autre candidat Y. Comme X est un vainqueur de Condorcet, il y a une majorité d'électeurs qui le préfèrent à Y. Cette majorité peut alors se concerter et se dire « votons tous pour X dès le premier tour ; ainsi il aura une majorité de suffrages sur son nom et sera élu, ce qui plus intéressant pour nous ! ». Bien sûr, cela ne signifie par pour autant que tous les électeurs de cette majorité aient X pour candidat préféré. On peut même imaginer qu'un nombre conséquent d'eux préfèrent en fait un troisième candidat Z, et puissent être tentés de faire gagner Z avec l'aide d'anciens électeurs de Y qui préfèrent également Z à X... Mais c'est peine perdue, car X étant un vainqueur de Condorcet, il y a *aussi* une majorité (mais pas constituée des mêmes électeurs, bien sûr) qui préfère X à Z ; si Z devient trop dangereux dans les sondages, c'est *cette* majorité qui se liguera alors pour voter X afin d'éviter l'élection de Z ! [8]

Ainsi, une méthode satisfaisant le critère de Condorcet a ceci d'avantageux que les électeurs n'ont pas à se creuser la cervelle pour savoir si leur vote est bien calculé : avec une telle méthode en effet, s'il y a un vainqueur de Condorcet, il sera nécessairement élu !

⁂

Une autre qualité des méthodes satisfaisant le critère de Condorcet est la suivante : alors que les méthodes classiques (notamment les méthodes A, B, C, D et E de l'article précédent) sont sujettes à un phénomène d'"éparpillement des voix" quand le nombre de candidats devient grand, phénomène qui peut conduire à un résultat non pertinent, le résultat d'une méthode satisfaisant le critère de Condorcet ne change pas quand on retire ou qu'on ajoute des "petits" candidats :

Théorème (Indifférence aux petits candidats)

> i. *Si on part d'une situation dans laquelle un des candidats est vainqueur de Condorcet, et qu'un ou plusieurs des autres candidats se retirent de l'élection, alors le vainqueur de Condorcet initial reste vainqueur de Condorcet après ce retrait.*
> ii. *Inversement, si on part d'une situation dans laquelle un candidat est vainqueur de Condorcet, et qu'on introduit un ou plusieurs nouveaux candidats à chacun desquels le vainqueur initial est préféré en face-à-face, alors le vainqueur de Condorcet initial reste vainqueur de Condorcet après cet ajout.*

Démonstration

> *Le retrait ou l'ajout d'un candidats ne change pas les préférences binaires des électeurs entre les autres candidats. Pour le point i, puisque notre vainqueur de Condorcet initial battait initialement tous ses adversaires en face-à-face, cela reste donc le cas si un ou plusieurs de ses adversaires se retirent. Pour le point ii, notre vainqueur de Condorcet initial continue de même à battre tous ses adversaires initiaux, et nous avons supposé qu'il battait également ses nouveaux adversaires, de sorte qu'il est bien toujours vainqueur de Condorcet dans la nouvelle situation.*

Notons en revanche que, si le vainqueur de Condorcet éventuel est dans un sens un vainqueur inéluctable, ce n'est par contre pas nécessairement celui qui est réellement le préféré du peuple (au sens que nous avions évoqué dans l'article précédent, c'est-à-dire en mesurant le bonheur total que ce candidat apporterait à la population s'il était élu). Ainsi, s'il n'y a que deux candidats X et Y, et que 51 % des électeurs préfèrent légèrement X à Y et 49 % des électeurs préfèrent très largement Y à X, c'est X qui sera élu alors qu'on a envie de dire que c'est plutôt Y qui représente la volonté du peuple. Cela dit, ce phénomène de « dictature de la majorité » [9] n'est pas directement lié au critère de Condorcet : c'est en fait un vice inhérent à la démocratie [10], comme l'avait déjà compris Kant [11].

⁂

Pour l'instant, nous avons beaucoup parlé de philosophie et peu de mathématiques (et encore moins de recherche !). Mais cette introduction était nécessaire pour comprendre pourquoi le critère de Condorcet est un objet mathématique naturel et important. Dans la seconde partie de cet article, nous allons maintenant parler de recherches beaucoup plus récentes qui tournent autour de ce critère.

Le paradoxe de Condorcet existe-t-il vraiment ?

Le paradoxe de Condorcet

Le critère de Condorcet nous amène naturellement à parler du *paradoxe de Condorcet*, que nous avions déjà rencontré dans **l'article précédent**. Ce paradoxe énonce en substance qu'il n'y a pas nécessairement de vainqueur de Condorcet :

Paradoxe de Condorcet (1785)

> *Dès qu'il y a au moins 3 candidats, il existe des situations où il n'y a pas de vainqueur de Condorcet.*

Démonstration

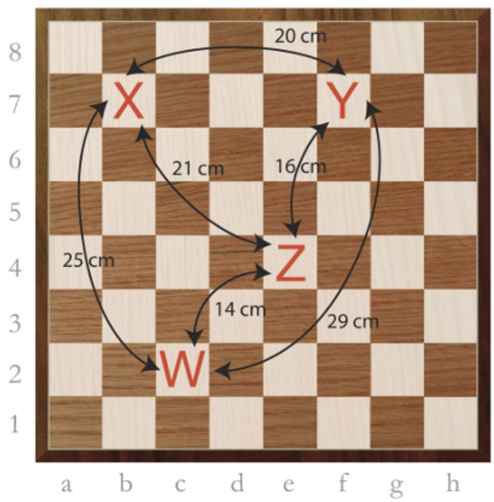
> *Supposons qu'il y ait au moins 3 candidats, et choisissons trois candidats particuliers parmi eux, appelés X, Y et Z. Nous imaginons que tous les électeurs préfèrent X, Y et Z à tous les autres candidats, de sorte qu'aucun de ces autres candidats ne peut être vainqueur de Condorcet. Maintenant, on peut imaginer en outre que 40 % d'électeurs ont l'ordre de préférence « X puis Y puis Z », 35 % « Y puis Z puis X » et 25 % « Z puis X puis Y ». Alors 65 % des électeurs préfèrent X à Y (donc Y n'est pas un vainqueur de Condorcet), 75 % préfèrent Y à Z (donc Z n'est pas un vainqueur de Condorcet), et 60 % préfèrent Z à X (donc X n'est pas un vainqueur de Condorcet) : aucun candidat n'est donc vainqueur de Condorcet.*

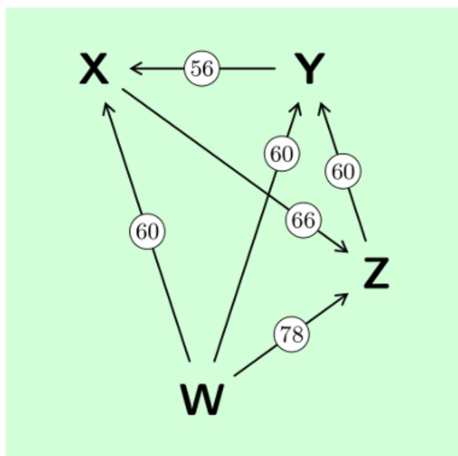
Toutefois, nous avons vu dans le préambule un cas d'élection réelle avec pas moins de 5 candidats dans lequel il existait bel et bien un vainqueur de Condorcet [12] ! Cela suggère qu'il pourrait éventuellement exister un phénomène qui interdirait le paradoxe de Condorcet en pratique, parce que les situations politiques réalistes vérifieraient certaines hypothèses particulières sous lesquelles le paradoxe est impossible. Dans les paragraphes suivants, nous allons présenter différentes modélisations des préférences politiques, et voir si celles-ci permettent ou pas l'occurrence du paradoxe de Condorcet.

Paradoxe de Condorcet sur un échiquier politique

Nous allons introduire ici une modélisation assez réaliste des préférences politiques des électeurs que nous appellerons *modélisation par échiquier politique*, dont l'idée est simplement qu'en général, nous votons pour le candidat dont les positions s'éloignent le moins de nos opinions. On supposera ainsi qu'il existe un « échiquier politique » abstrait (qui permet de parler de « distance politique ») sur lequel on peut placer aussi bien les candidats que les électeurs, tel que les préférences de chaque électeur aillent du candidat qui lui est le plus proche vers celui qui lui est le plus éloigné.

On pourrait espérer qu'avec une telle modélisation, il n'y ait jamais de paradoxe de Condorcet... mais ce n'est pas le cas ! Supposons en effet que l'échiquier politique corresponde à un véritable échiquier (dont les cases mesurent, disons, 5 cm de côté), où l'on place un candidat X sur la case b7, un candidat Y sur la case f7, un candidat Z en e4 et un candidat W en c2. Nous imaginons que l'électorat est divisé en quatre groupes de tailles diverses dont chacun a exactement la même position qu'un des candidats [13] : 34 % des électeurs sont au niveau de X, 26 % au niveau de Y, 18 % de Z et 22 % de W. Les distances entre candidats sont alors celles qui sont indiquées sur le dessin ci-dessous, de sorte qu'on a les préférences suivantes pour les électeurs (« X>Y » signifiant « X est préféré à Y ») :





L'exemple de l'échiquier

Voici, à gauche, une configuration d'échiquier politique telle que, s'il y a 34 % des électeurs en X, 26 % en Y, 22 % en Z et 18 % en W, et que chaque électeur préfère les candidats dont il est le plus proche, alors aucun candidat n'est vainqueur de Condorcet : à droite, le diagramme de préférences binaires correspondant.

- X > Y > Z > W pour 34 % des électeurs ;
- Y > Z > X > W pour 26 % des électeurs ;
- W > Z > X > Y pour 22 % des électeurs ;
- Z > W > Y > X pour 18 % des électeurs.

Alors, dans cette configuration, X n'est pas un vainqueur de Condorcet (parce qu'il est battu par Z à 66 %), ni Y (battu par X à 56 %), ni Z (battu par Y à 60 %), ni W (battu par Z à 74 %) : on a donc bien paradoxe de Condorcet !

Les théorèmes de l'électeur médian

Dans l'exemple de l'échiquier, nous avons fabriqué un échiquier politique *de dimension 2*, c'est-à-dire avec 2 coordonnées (une pour les colonnes et une pour les lignes). Ce n'était pas un hasard ! Car si l'échiquier politique n'a qu'une seule dimension, on peut démontrer qu'il existe toujours un vainqueur de Condorcet :

Théorème (Premier théorème de l'électeur médian ; Black 1958)

> *Lorsque l'échiquier politique est un axe gradué (qu'on interprète généralement comme un positionnement politique du "plus à gauche" vers le "plus à droite"), il y a toujours un vainqueur de Condorcet.*

Pour démontrer ce théorème, nous allons d'abord introduire une définition qui nous sera utile dans la preuve :

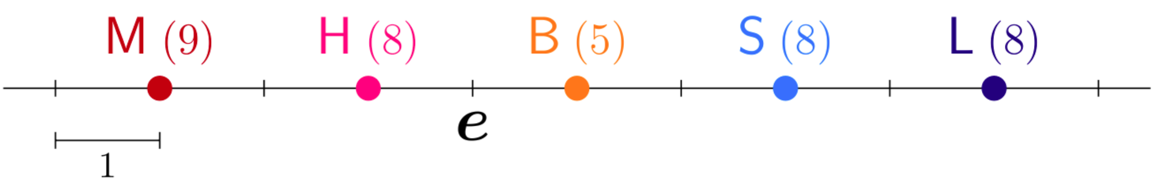
Définition (Profil de préférences en Λ)

> *On dit qu'un électeur a un profil de préférences en Λ (lire « lambda ») quand, lorsqu'on parcourt la liste des candidats du plus à gauche au plus à droite, la préférence de cet électeur monte jusqu'à atteindre son candidat préféré, puis redescend. (Éventuellement, le candidat préféré peut être le plus à gauche ou le plus à droite de tous ; l'important étant de ne jamais remonter après être descendu).*
>
> *Par exemple, si nous supposons que les candidats de l'élection présidentielle se rangent de gauche à droite selon l'ordre « M, H, B, S, L », alors les préférences « B>H>S>L>M » ou « L>S>B>H>M » sont des profils en Λ, mais pas « H>M>S>B>L », puisqu'une telle préférence baisse entre H et B avant de remonter entre B et S.*

Démonstration du premier théorème de l'électeur médian

> *Le premier point est de constater que, quand l'échiquier politique est un axe gradué,* ***tous les électeurs ont un profil de préférence en Λ****. En effet, considérons n'importe quel électeur. Quand on parcourt l'échiquier politique de gauche à droite, on commence par se rapprocher de cet électeur avant de s'en éloigner, de sorte que les préférences de l'électeur augmentent (tant qu'on est plus à gauche que cet électeur) avant de redescendre (une fois qu'on est plus à droite) : cela correspond bien à un profil de préférences en Λ. (Si aucun candidat n'est situé exactement à la même position que l'électeur, il faudra regarder les distances précises pour savoir si le candidat préféré de cet électeur est celui situé immédiatement à sa gauche ou immédiatement à sa droite, mais dans les deux cas on a bien un profil en Λ).*
>
> *Nous prouvons maintenant le théorème. Pour éviter les problèmes d'ex æquo, supposons qu'il y a un nombre impair d'électeurs. Chacun de ces électeurs a un candidat préféré. Imaginons que nous demandions à chaque électeur d'écrire le nom de son candidat préféré sur un bulletin puis que nous classions les bulletins obtenus depuis le candidat le plus à gauche vers le candidat le plus à droite, et appelons X le candidat qui est sur le bulletin situé au milieu de ce classement. J'affirme alors que X est un vainqueur de Condorcet. En effet, considérons un candidat Y situé, par exemple, plus à droite que X (pour un candidat plus à gauche, il suffirait d'inverser « droite » et « gauche » dans la suite). Comme X est au milieu de notre classement des bulletins, il y a une majorité absolue de bulletins qui sont au nom soit de X, soit d'un candidat plus à gauche que X, ce qui signifie qu'il y a une majorité absolue d'électeurs dont le candidat préféré est soit X, soit un candidat situé plus à gauche que X. Considérons alors un tel électeur, dont nous appelons Z le candidat préféré (Z pouvant éventuellement être le même que X). Nous savons par le paragraphe ci-dessus que notre électeur a un profil de préférences en Λ, donc quand on s'éloigne de Z vers la droite, les préférences de l'électeur ne font que décroître. Or ici X est au moins aussi à droite que Z et Y est encore plus à droite que X, donc on croisera Y postérieurement à X au cours de cet éloignement ; par conséquent notre électeur préfère X à Y. Comme il en est de même pour une majorité absolue d'électeurs, cela signifie que X est majoritairement préféré à Y, et comme on peut faire le même raisonnement pour tous les Y possibles, cela prouve que X est un vainqueur de Condorcet.*
>
> *Notez que cette démonstration montre que le théorème reste valable dès que tous les électeurs ont un profil de préférences en Λ.*

Le défaut de ce premier théorème de l'électeur médian (et de la modélisation par échiquier politique en général) est qu'il considère que la seule qualité des candidats qui intéresserait les électeurs est leur positionnement politique. Or, à positionnement politique égal, il se peut que certains candidats soient "meilleurs" ou "moins bons", par exemple à cause de leurs qualités intellectuelles ou charismatiques. Un modèle plus fin est donc le suivant : d'une part, il y a un axe gradué sur lequel sont placés les candidats et les électeurs, et d'autre part est aussi associée à chaque candidat une « valeur intrinsèque ». Chaque électeur calcule alors pour les différents candidats un « score » égal à la valeur intrinsèque du candidat moins la distance du candidat, et classe les candidats par score décroissant : un exemple est donné sur le dessin ci-dessous.



Axe politique avec valeurs intrinsèques

Voici un exemple tenant compte à la fois du placement des candidats sur l'axe politique « gauche – droite » et de la valeur intrinsèque des candidats. Dans cet exemple, il y a cinq candidats repérés par des points, la valeur intrinsèque de chacun d'eux étant indiquée entre parenthèses. Si nous considérons un électeur situé à la position « e », cet électeur donne alors à M un score égal à 9 (valeur) – 5 (distance) = 4, à H un score de 8 – 1 = 7, à B un score de 5 – 1 = 4, et de même à S et L des scores respectifs de 5 et 3. Par conséquent, l'ordre de préférences de cet électeur est « H > M > S > B > L » — qui n'est pas un profil en Λ.

Eh bien, dans ce modèle, on est encore assuré de l'existence d'un vainqueur de Condorcet :

Théorème (Second théorème de l'électeur médian ; Roberts 1977)

> *Si les candidats et les électeurs sont placés sur un même axe gradué d'une part, qu'à chaque candidat est associée une valeur intrinsèque d'autre part, et que les préférences de chaque électeur sont déterminées par la différence « valeur intrinsèque − distance du candidat », alors il y a toujours un vainqueur de Condorcet.*

Pour démontrer ce théorème, nous commençons ici aussi par introduire une définition :

Définition (Préférences décroisées)

> *Étant donnés des candidats rangés de gauche à droite et des électeurs également rangés de gauche à droite, on dit que la configuration des préférences est décroisée s'il est impossible de trouver un candidat X plus à gauche qu'un autre candidat Y et un électeur A plus à gauche qu'un autre électeur B qui soient tels que A préfère Y à X mais que B préfère X à Y (autrement dit, qui soient tels que l'électeur de gauche préfère le candidat de droite et l'électeur de droite le candidat de gauche).*

Démonstration du second théorème de l'électeur médian

Le premier point est de constater que, sous les hypothèses du théorème, **la configuration des préférences est décroisée***. Considérons en effet deux candidats X et Y, avec Y plus à droite que X. Dire que la configuration des préférences est décroisée signifie que, à mesure qu'un électeur fictif se déplace de gauche à droite sur l'axe politique, il ne peut pas passer d'une préférence « Y>X » à une préférence « X>Y » ; pour montrer cela, il suffit donc d'établir que la différence « score de Y − score de X » ne peut qu'augmenter quand on va de gauche à droite. Or, dans cette différence de scores, la seule chose qui change d'un électeur à l'autre est la partie du score qui dépend de la distance : celle qui dépend du score est la même ! Ainsi, il nous suffit en fait de montrer que la différence « distance de Y − distance de X » ne fait que diminuer quand on va de la gauche vers la droite. Or donc, que se passe-t-il quand on va de la gauche vers la droite ? Dans un premier temps, tant qu'on est encore à gauche de X, chaque pas vers la droite nous rapproche autant de X que de Y, et donc la différence « distance à Y − distance à X » ne varie pas. Puis, quand on se trouve entre X et Y, chaque pas vers la droite fait non seulement diminuer la distance à Y, mais aussi augmenter celle à X, de sorte que la différence « distance à Y − distance à X » diminue. Enfin, au-delà de Y, chaque pas vers la droite nous éloigne autant de X que de Y, et donc à nouveau la différence des distances reste constante. Ainsi on a bien une configuration de préférences décroisée.*

*Nous prouvons maintenant le théorème. Pour éviter les problèmes d'ex-æquo, supposons qu'il y a un nombre impair d'électeurs. Si nous classons ces électeurs de gauche à droite sur l'axe politique, il y en a alors un qui se retrouve au milieu de ce classement, que nous appellerons l'« électeur médian ». J'affirme que les préférences majoritaires sont exactement les mêmes que celles de cet électeur médian, et donc qu'en particulier le candidat préféré de l'électeur médian est un vainqueur de Condorcet [***14***]. En effet, supposons qu'entre deux candidats X et Y, l'électeur médian préfère X ; pour fixer les idées, disons que Y est plus à droite que X (sinon, inverser « droite » et « gauche »). Alors d'après le résultat intermédiaire ci-dessus, tous les électeurs plus à gauche que l'électeur médian préfèrent également X à Y, et donc il y a une majorité d'électeurs qui préfèrent X à Y, ce qu'on voulait démontrer.*

*Notez que cette démonstration montre que le théorème reste valable dès que la configuration des préférences est décroisée [***15***].*

En plus de démontrer l'existence du vainqueur de Condorcet, la preuve ci-dessus nous dit à quoi ce vainqueur correspond : il s'agit du candidat préféré de l'électeur situé "au milieu" de tous sur l'axe des préférences (d'où le nom de « théorème de l'électeur médian »). Un phénomène similaire se produisait pour le premier théorème de l'électeur médian, où le résultat de l'élection correspondait au choix le plus central parmi les candidats favoris des différents électeurs. Une moralité qu'on peut en tirer est que le critère de Condorcet favorise les candidats "centristes" par rapport aux candidats plus orientés ou extrémistes. Est-ce un comportement souhaitable ? Les avis divergent sur cette question [**16**] : pour certains, favoriser les centristes risque de dégoûter les électeurs non centristes qui auraient l'impression de ne jamais être pris en compte, et de favoriser la politique du consensus mou au détriment de projets plus ambitieux ; d'autres, au contraire, estiment qu'une politique centriste apaiserait le débat en permettant à chacun de s'y retrouver un peu, et qu'une grande stabilité politique serait plus propice à la mise en place de projets à long terme...

Et quand il n'y a pas de vainqueur de Condorcet ?

Les théorèmes de l'électeur médian tendent à indiquer que dans les situations concrètes, on aura la plupart du temps un vainqueur de Condorcet, ce que corrobore effectivement l'histoire des élections [**17**]. Cela dit, la modélisation purement unidimensionnelle sur laquelle reposent ces théorèmes est clairement trop réductrice, et l'exemple de l'échiquier nous montre qu'il suffit d'un cas à peine plus complexe pour voir réapparaître le paradoxe de Condorcet... Bref, si on souhaite appliquer une méthode électorale satisfaisant le critère de Condorcet, il faudra donner une règle pour désigner le candidat élu quand il n'y a pas de vainqueur de Condorcet. Mais laquelle ? Cette question n'a commencé à être étudiée qu'il y a quelques décennies [**18**], notamment parce qu'avant l'invention de l'informatique il était pratiquement impossible de réaliser le dépouillement d'un vote où les électeurs indiquent leurs préférences entre *tous* les candidats. C'est des recherches sur ce sujet que nous allons maintenant parler.

La méthode minimax

Commençons par une approche très simple. On peut se dire : « à la base, un vainqueur de Condorcet est un candidat qui réalise plus de 50 % face à tous les autres. Si aucun des candidats n'y parvient, on peut baisser le seuil des 50 % pour dire : "s'il y a un seul des candidats qui parvient à atteindre, par exemple, 45 % face à n'importe quel autre, alors c'est ce candidat qui doit être élu" ». Autrement dit, le protocole est le suivant : pour chacun des candidats, on regarde le pire des scores qu'il ferait dans ses différents faces-à-faces, puis on déclare élu celui des candidats dont le pire des scores est le meilleur. Cette méthode est appelée *méthode minimax*, car elle regarde qui est le meilleur (« max ») dans la situation où il est le moins bon (« mini »).

Ainsi, pour l'exemple de l'échiquier présenté plus haut (où il n'y a pas de vainqueur de Condorcet), la pire défaite de X a lieu contre Z par 66 %, la pire défaite de Y contre X par 56 %, la pire défaite de Z contre Y par 60 % et la pire défaite de W contre Z par 78 %. La plus légère de ces pires défaites est celle de Y, qui est donc le vainqueur minimax.

Cela semble tout naturel ! Hélas, cette méthode ouvre la voie à une forme de "manipulation" (synonyme péjoratif de « vote stratégique ») particulièrement dangereuse. Supposons en effet que dans l'exemple de l'échiquier abordé précédemment, nous ayons en plus un groupe V composé d'"anarchistes" [**19**]. Ces anarchistes sont relativement nombreux puisqu'ils constituent 40 % de la population totale, de sorte qu'on a la distribution globale suivante : X 20 %, Y 16 %, Z 11 %, W 13 %, V 40 %. Cependant les anarchistes font l'unanimité contre eux en-dehors de leur groupe : aussi bien les partisans de X que de Y que de Z que de W classent V en dernière position dans leur ordre de préférences ! C'est là que les anarchistes ont une idée géniale pour prendre le pouvoir malgré cet important handicap. Ils se disent : « plutôt que de voter de façon équilibrée entre nos concurrents, nous allons nous arranger pour amplifier le plus possible les pires défaites de chacun d'eux ! ».

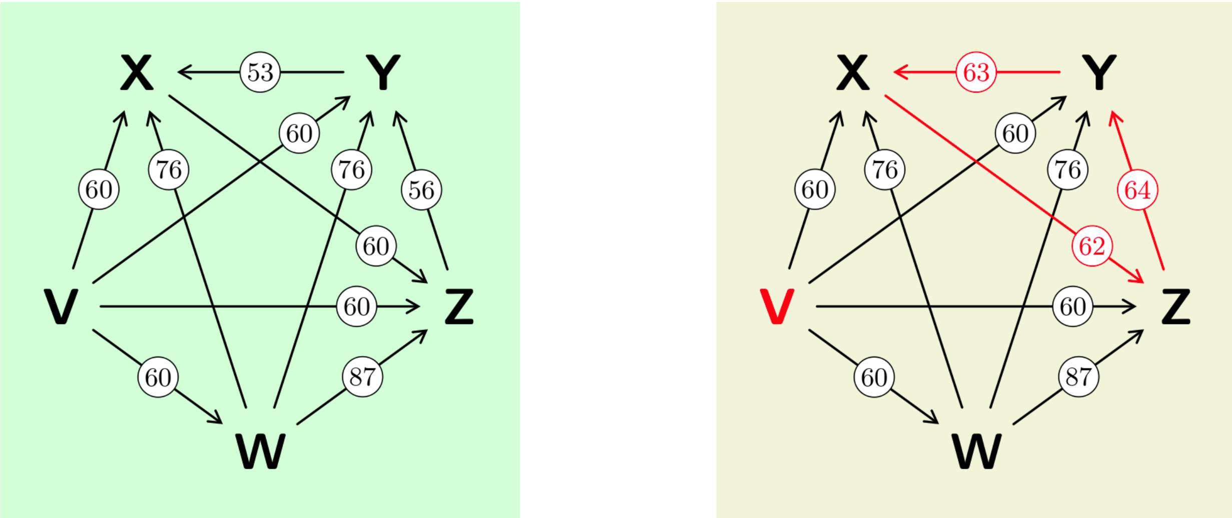
Pour ce faire, les électeurs du groupe V se divisent en trois sous-groupes (par contre, ils ne présentent toujours qu'un seul candidat) :

- Le sous-groupe V_1 compte 18 % de tous les électeurs et vote « V > Z > Y > X > W » ;
- Le sous-groupe V_2 compte 12 % des électeurs et vote « V > Y > X > Z > W » ;
- Le sous-groupe V_3 compte 10 % des électeurs et vote « V > X > Z > Y > W ».

Avec cette stratégie, on a donc les votes suivants :

- X > Y > Z > W > V pour 20 % des électeurs ;
- V > X > Y > Z > W pour 18 % des électeurs ;
- Y > Z > X > W > V pour 16 % des électeurs ;
- W > Z > X > Y > V pour 13 % des électeurs ;
- V > Z > X > Y > W pour 12 % des électeurs ;
- Z > W > Y > X > V pour 11 % des électeurs ;
- V > Y > Z > X > W pour 10 % des électeurs.

Cela conduit au diagramme de préférences binaires suivant (à droite) :



La manipulation des anarchistes

Voici les diagrammes correspondant à l'exemple de l'échiquier avec anarchistes, à gauche dans le cas où les anarchistes votent sincèrement (on suppose que les préférences des anarchistes entre X, Y et Z sont aléatoires et que W est toujours leur candidat détesté), et à droite avec la manipulation que nous avons expliquée (les changements sont indiqués en rouge). Dans le premier cas, le vainqueur minimax est Y ; dans le second, c'est V, à qui sont pourtant préférés tous les autres candidats !

Alors, les pires défaites de X, Y et Z sont respectivement de 62 %, 63 % et 64 % (respectivement contre Z, X et Y), la pire défaite de W est de 87 % (contre Z) et la pire défaite de V est de 60 % (contre n'importe lequel de ses adversaires) : c'est donc V qui devient le vainqueur minimax !

Évidemment, les partisans de X, Y , Z et W ont alors la possibilité de s'allier pour désigner Y comme candidat commun et faire échouer le plan de V, mais nous avons dit quelques paragraphes plus haut que la philosophie du critère de Condorcet était d'épargner autant que possible aux électeurs les tourments du vote stratégique ! Il semble donc, selon ce point de vue, que la méthode minimax ne soit pas pertinente quand il n'y a pas de vainqueur de Condorcet.

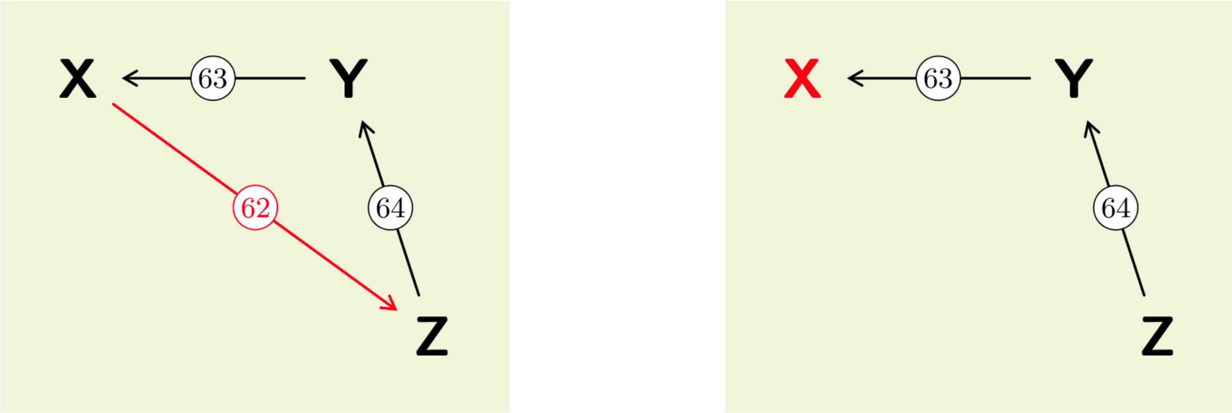
La méthode Schulze

Nous allons maintenant présenter une méthode conçue en 1997 par un étudiant en physique allemand du nom de Markus Schulze, qui ne présente pas le handicap qu'avait la méthode minimax. La philosophie derrière la méthode Schulze est la suivante : « au fond, une élection, ce n'est jamais qu'une astuce pour éviter les désagréments des révolutions : on consulte les électeurs afin de savoir qui finirait par gagner la révolution sans avoir à se battre ! ».

Ce qu'on appellera une « révolution », ici, c'est quand un grand nombre d'électeurs se mettent d'accord pour faire tomber le président et le remplacer par un autre, sur l'identité duquel ils se sont entendus préalablement. Dans ce sens-là, il est facile de comprendre le critère de Condorcet : si un candidat (appelons-le X) est vainqueur de Condorcet, n'importe quel autre candidat qui arriverait au pouvoir serait rapidement renversé par la majorité qui préfère X à ce candidat, alors qu'une fois que X sera au pouvoir il n'y aura jamais de majorité pour le renverser, ce qui rendra son assise beaucoup plus stable [**20**]. Au final, puisqu'on a rapidement des transitions de n'importe qui vers X et rarement des transitions de X vers n'importe qui, c'est bien X qui sera au pouvoir la plupart du temps. (Notez bien qu'ici nous ne prétendons pas qu'il est impossible de renverser un président sans avoir une majorité liguée contre lui : nous disons juste que cela est très nettement plus facile de le faire dès qu'on a un peu plus de personnes pour mener la révolution).

Regardons alors ce qui arriverait dans un tel cadre pour l'exemple de l'échiquier avec anarchistes. Les révolutions ont tendance à renverser les présidents dans le sens indiqué par les flèches, et beaucoup plus rarement en sens inverse. Par conséquent, au bout de quelques révolutions, le président est forcément X, Y ou Z, puisqu'une fois qu'on est arrivé à un de ces trois-là il n'y a aucun chemin suivant les flèches qui ramène en W ou V ! On voit donc déjà (comme cela nous paraissait naturel) que ce n'est pas V que les révolutions amèneront au pouvoir la plupart du temps.

Poursuivons. Nous savons que ce seront essentiellement X, Y et Z qui se partageront le pouvoir, mais dans quelles proportions ? Si nous regardons juste les flèches concernant ces trois candidats, nous avons le diagramme de gauche ci-dessous :



La méthode Schulze contre la manipulation des anarchistes

Ces diagrammes expliquent ce qui se passe quand on applique la méthode Schulze à la situation où les anarchistes tentent leur manipulation. Le groupe de tête du diagramme des préférences binaires dans cette situation [voir la figure précédente] est alors constitué des trois candidats {X, Y, Z} : en particulier V est éliminé puisqu'il était battu par tous les autres candidats ! Dans un second temps, on regarde le diagramme des préférences binaires entre ces trois candidats [à gauche]. Pour faire apparaître un groupe de tête plus petit, on supprime la flèche la plus faible [en rouge]. Le nouveau diagramme [à droite] a alors un groupe de tête constitué du seul candidat X : c'est donc lui qui est déclaré vainqueur par la méthode.

Maintenant, l'idée est que, de même qu'il était beaucoup plus difficile de faire une révolution avec moins de 50 % du peuple qu'avec plus de 50 %, il est beaucoup plus difficile de faire une révolution avec 62 % qu'avec 63 % ou 64 %... Du coup, on efface la flèche marquée « 62 % » pour signifier que cette révolution est rare, et on arrive au diagramme de droite ci-dessus. Au bout de quelques révolutions, c'est alors forcément X qui est président : l'idée de la méthode Schulze est donc de dire que c'est lui qui doit être proclamé vainqueur de l'élection [21] !

⁂

Maintenant que nous avons expliqué le principe de la méthode Schulze, nous allons pour finir en donner une définition rigoureuse et nous assurer que celle-ci donne bien un vainqueur unique dans tous les cas. Nous supposerons simplement qu'il n'y a jamais deux candidats exactement à égalité ni deux flèches ayant exactement la même force dans le diagramme des préférences binaires, ce qui est toujours le cas en pratique quand il y a beaucoup d'électeurs.

Notez que dans ce paragraphe, nous rencontrerons parfois des diagrammes de préférences binaires *incomplets*, c'est-à-dire où certains candidats ne sont pas reliés par des flèches.

Définition (Groupe de tête)

> *Dans un diagramme de préférences binaires (éventuellement incomplet), on dit qu'un candidat (appelé ici X) appartient au groupe de tête quand, en partant de n'importe quel autre candidat, il est possible de suivre un chemin selon les flèches qui part de ce candidat et aboutit en X.*

Ici, il est important de remarquer que si le diagramme est complet, le groupe de tête n'est jamais vide :

Théorème

> *Dans un diagramme de préférences binaires complet, le groupe de tête n'est jamais vide.*

Démonstration

*Nous allons présenter une façon de construire le groupe de tête qui montre que celui-ci n'est jamais vide. J'appellerai ici « surgroupe de tête » tout groupe de candidats tel que, quand on part de n'importe quel candidat **qui n'est pas dans ce groupe**, on peut arriver à n'importe quel candidat du groupe en suivant les flèches.*

J'observe d'abord que le groupe de tous les candidats est un surgroupe de tête, puisqu'il n'y a aucun candidat qui n'appartienne pas à ce groupe !

Maintenant, je dis la chose suivante : quand on a un surgroupe de tête,

- *Ou bien ce surgroupe de tête est le vrai groupe de tête ;*
- *Ou bien, parmi ce surgroupe de tête, on peut sélectionner un certain nombre **non nul** de candidats qui forment un surgroupe de tête **strictement** plus petit.*

*En effet, supposons que notre surgroupe de tête ne soit pas le vrai groupe de tête. La seule façon que cela arrive est qu'il existe au moins un candidat X dans ce surgroupe de tête auquel il ne soit pas possible d'aboutir en partant d'un autre candidat Y, **lequel est nécessairement dans le surgroupe de tête**, puisqu'on sait qu'on peut aboutir à n'importe quel candidat du surgroupe de tête à partir de n'importe quel candidat hors du surgroupe de tête. Maintenant, je sélectionne les candidats du surgroupe de tête qui sont tels que, quand on part de ces candidats, on ne peut pas arriver en X. Ce nouveau groupe n'est pas vide puisqu'il contient Y, et il est strictement plus petit que le surgroupe de tête initial puisqu'il ne contient pas X. J'affirme que ce nouveau groupe est encore un surgroupe de tête. En effet, considérons un candidat Z qui n'est pas dans ce nouveau groupe. Soit ce candidat n'était pas dans notre surgroupe de tête initial, et alors sait qu'on peut aller de Z à n'importe quel point de notre surgroupe de tête initial et donc en particulier à n'importe quel point de notre nouveau groupe. Soit ce candidat était dans notre surgroupe de tête initial mais pas dans le nouveau groupe. Alors, du candidat Z, on peut revenir au candidat X en suivant les flèches (par définition du nouveau groupe). Mais du candidat X, on peut aller vers n'importe quel candidat W du nouveau groupe : en effet, il existe une flèche entre X et W vu que notre diagramme est complet, et celle-ci pointe nécessairement de X vers W puisqu'il n'y a pas de chemin de W vers X ! Au final, on a montré que pour tout candidat Z qui n'est pas dans notre nouveau groupe, on peut aller de ce candidat à n'importe quel candidat du nouveau groupe en suivant les flèches, ce qui signifie bien que le nouveau groupe est un surgroupe de tête.*

Partons maintenant du surgroupe de tête formé par tous les candidats, et tant qu'on n'a pas encore trouvé le vrai groupe de tête, appliquons la méthode ci-dessus pour former des surgroupes de tête de plus en plus petits, mais tous non vides. Forcément à un moment on sera obligé de s'arrêter, et à ce moment-là notre surgroupe de tête sera le vrai groupe de tête d'après l'alternative ci-dessus, et ce groupe de tête sera bien non vide, ce qui prouve le théorème.

Définition (Méthode Schulze)

> *À partir d'un diagramme de préférences binaires complet (où toutes les flèches portent des scores différents), on définit le vainqueur de Schulze par la méthode suivante :*

> i. *On part du diagramme des préférences binaires pour tous les candidats.*
> ii. *S'il n'y a en fait qu'un seul candidat, celui-ci est proclamé vainqueur et on s'arrête.*
> iii. *Sinon, on regarde le groupe de tête : s'il strictement plus petit que l'ensemble initial, on efface tous les candidats qui ne sont pas dans ce groupe de tête (ainsi que les flèches les concernant), et on recommence la méthode en faisant comme s'il n'y avait plus les candidats restants.*
> iv. *Si, en revanche, le groupe de tête contient tous les candidats, on efface progressivement les flèches de plus petits scores, jusqu'à ce que le groupe de tête du diagramme obtenu en effaçant les flèches exclue au moins un des candidats (ce qui arrive forcément, car une fois que toutes les flèches seraient effacées il n'y aurait plus personne dans le groupe de tête). À ce moment-là, on efface tous les candidats qui ne sont pas dans ce nouveau groupe de tête (ainsi que les flèches les concernant), et on recommence la méthode en faisant comme s'il n'y avait plus que les candidats restants.*

Pour s'assurer que la méthode conduit bien à la désignation d'un vainqueur, il reste à vérifier que, au cours de la procédure d'effacement des flèches, on ne risque pas de tomber sur un groupe de tête vide :

Théorème

> *Si on a un diagramme de préférences binaires pour plusieurs candidats (éventuellement incomplet) dans lequel le groupe de tête contient tous les candidats, alors effacer une seule des flèches ne peut pas amener à une situation où le nouveau groupe de tête serait vide. Par conséquent, quand on part d'un diagramme de préférences binaires dans lequel le groupe de tête contient tous les candidats et qu'on efface ses flèches une à une, le premier moment où le groupe de tête cesse de contenir tous les candidats donne un groupe de tête non vide.*

Démonstration

> *Nous ne démontrons que le début de l'énoncé, la partie qui suit « par conséquent » en découlant immédiatement. Partons donc d'un diagramme de préférences binaires dans lequel le groupe de tête contient tous les candidats, et effaçons-en une flèche, dont nous appelons X le candidat duquel elle partait. Nous allons montrer que X est forcément dans le nouveau groupe de tête, qui sera donc non vide. Pour ce faire, considérons n'importe quel candidat Y, et montrons qu'on peut aller de Y à X en suivant les flèches malgré l'effacement effectué. Comme, avant l'effacement de la flèche, le groupe de tête contenait tous les candidats, on avait alors un chemin qui allait de Y à X. Mais forcément ce chemin n'empruntait pas la flèche effacée, puisque cette flèche part de X alors que dans notre chemin X est le point d'arrivée ! C'est donc que le chemin peut toujours être suivi une fois la flèche effacée, ce qu'on voulait.*

⁂

Nous avions dit que l'intérêt de la méthode Schulze était d'éviter les manipulations comme celle des anarchistes dans l'exemple de l'échiquier. On peut effectivement prouver qu'une telle manipulation est impossible avec cette méthode :

Théorème (Robustesse à la manipulation)

> *Supposons qu'il y a deux types de candidats : les candidats « déraisonnables » et les candidats « raisonnables » (on suppose qu'il y a au moins un candidat raisonnable), et appelons « électeur raisonnable » un électeur qui classe tous les candidats déraisonnables à la fin de son ordre de préférences. Alors, s'il y a une majorité absolue d'électeurs raisonnables et que ces électeurs votent sincèrement, aucun candidat déraisonnable ne peut gagner avec la méthode Schulze.*

Démonstration

> *D'après la définition de la méthode Schulze, il suffit de montrer qu'aucun candidat déraisonnable n'appartient au groupe de tête dans une telle situation. Or, quand on part d'un candidat raisonnable et qu'on suit une flèche du diagramme des préférences binaires, on aboutit forcément à un autre candidat raisonnable : en effet, la majorité d'électeurs raisonnables assure qu'aucune flèche ne peut pointer d'un candidat raisonnable vers un candidat déraisonnable. Par conséquent, il est impossible d'avoir un chemin qui parte d'un candidat raisonnable et aboutisse à un candidat déraisonnable, ce qu'on voulait.*

Conclusion

Le critère de Condorcet pour une méthode électorale démocratique, qui énonce qu'un candidat préféré à n'importe quel autre en face-à-face doit toujours être élu, permet dans une certaine mesure d'éviter les problèmes du vote stratégique. Quand il n'y a pas de vainqueur de Condorcet, on peut imaginer plusieurs méthodes « de Condorcet » différentes : parmi elles, la méthode Schulze a l'avantage d'avoir à la fois une justification heuristique simple et d'être plus robuste à certaines formes de manipulation que, par exemple, la méthode minimax. C'est actuellement une des méthodes de vote les plus prisées par les théoriciens de la démocratie, et c'est elle qui est utilisée par exemple par les développeurs du système d'exploitation libre *Debian* — les geeks étant toujours à la pointe des progrès bizarres... ;-)

Dans un prochain article, nous présenterons d'autres méthodes originales, ne satisfaisant pas forcément le critère de Condorcet, mais qui ont également été préconisées par les théoriciens pour leurs propriétés mathématiques.

Références

- En piste bleue : les pages *Wikipédia* **« Méthode Condorcet »**, **« Paradoxe de Condorcet »** et **« Méthode Schulze »** (cette dernière page étant plutôt en piste rouge).
- Sur les théorèmes de l'électeur médian : *Fundamentals of Social Choice Theory*, par Roger B. Myerson (1996) ; texte librement disponible **sur l'internet** (hors-piste).
- **wiki.electorama.com** : Un wiki qui présente notamment les différentes méthodes de Condorcet et leurs propriétés respectives (hors-piste).

P.S. :

Je tiens à remercier chaleureusement la Rédaction du site, qui, afin que cet article puisse paraître en phase avec l'actualité, n'a pas hésité à prendre en charge elle-même la relecture du texte : merci donc aux relecteurs internes (notamment Julien Melleray et Clément Caubel) pour leurs conseils judicieux, et aux membres du Comité de Rédaction qui ont mis en place cette procédure exceptionnelle. Merci aussi à ma sœur « Ἀνεπώνυμη » qui a gentiment accepté de sous-traiter la réalisation du dessin de l'échiquier, ainsi qu'à mes collègues Aline K., Anne B. et Olivier G. pour leur assistance dans la traque aux coquilles !

Notes

[1] L'auteur pris en compte les voix des "petits" candidats en reportant leurs voix sur les "gros" selon les proximités politiques généralement admises.

[2] En fait, certains de ces faces-à-faces alternatifs ont réellement été testés par les sondeurs ; cependant la règlementation française interdit de publier un sondage de second tour dont l'affiche ne correspondrait pas à celle prédite par les sondages de premier tour [voir **le site de la Commission des Sondages**, paragraphe « Cohérence des questions »].

[3] Ces chiffres ont été obtenus par l'auteur à l'aide d'un modèle mathématique pour reconstituer les données manquantes à partir des résultats de sondages dont il disposait. Ces "résultats" ne prétendent donc pas refléter fidèlement la réalité et sont à prendre comme une simple hypothèse de travail.

[4] NOTA – L'auteur tient à préciser que la question est à prendre dans un sens strictement rhétorique : cet article n'a aucune visée partisane et le but de son introduction est purement pédagogique.

[5] Compilation d'extraits du discours préliminaire de l'*Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*.

[6] La justification de Condorcet lui-même n'était en fait pas fondée sur la notion de *volonté du peuple* mais plutôt sur celle de *vérité* : Condorcet considérait les électeurs comme autant de juges devant évaluer quel est le meilleur candidat et interprétait les divergences d'opinions comme des erreurs d'appréciation malencontreuses de ces juges. Ici nous suivrons une approche foncièrement différente.

[7] Cela est également valable dans un grand nombre de méthodes de vote, comme les méthodes A et F de notre article précédent. Par contre, ce n'est pas valable dans la méthode C (où les bulletins représentent le candidat *contre* lequel on vote et où c'est celui qui a le *moins* de voix qui gagne). Dans cette méthode en effet, si X est un candidat classé en tête par 60 % des électeurs (et donc en particulier vainqueur de Condorcet) et que Y et Z sont deux candidats de positions très proches classés en tête ensemble par 40 % des électeurs, il se passera la chose suivante : les partisans de Y et Z donneront 40 % de suffrages contre X, tandis que les partisans de X ne seront pas assez nombreux pour faire éliminer *à la fois* Y et Z, puisque cela demanderait que chacun de ces deux candidats reçoive 40 % de suffrages contre lui, soit 80 % en tout : X ne pourra donc pas être élu malgré la majorité d'électeurs en sa faveur !

[8] Là, vous m'objecterez : « mais puisque, selon vous, le vainqueur de Condorcet peut toujours s'arranger pour gagner, comment se fait-il que Bayrou ait été éliminé au premier tour de l'élection présidentielle ? ». Eh oui, c'est là toute la différence entre la théorie et la pratique... On peut avancer trois explications pour ce phénomène, qui ont sans doute joué toutes les trois :

- Première explication : les électeurs ont un haut sens moral et se refusent à voter de façon "stratégique".
- Seconde explication : le mécanisme de vote stratégique que nous avons présenté exige que les électeurs préférant Bayrou à Hollande (par exemple) se concertent pour changer leur vote *tous ensemble*, faute de quoi l'effet risque de ne pas être celui prévu ! Ce qui est irréalisable quand il s'agit d'une concertation entre des millions d'électeurs.
- Troisième explication : le mécanisme de vote stratégique exige aussi, pour prendre en compte de façon pertinente les intentions de vote du reste de la population, que les électeurs disposent d'une information parfaitement fiable sur celles-ci. Or en pratique, nous savons bien que les chiffres des instituts de sondages souffrent d'une certaine erreur... En outre, les électeurs espèrent souvent que la dynamique de campagne changera les intentions de vote en faveur de leur candidat chéri, et peuvent donc préférer prendre le risque que gagne un candidat qu'ils détestent pour garder une chance que leur candidat favori l'emporte.

[9] L'expression est due à **Tocqueville**.

[10] Du moins à partir du moment où on suppose que les électeurs cherchent d'abord à servir leur intérêt personnel.

[11] Le philosophe écrit ainsi dans l'article I de son opuscule *Vers la paix perpétuelle* : « La Démocratie est nécessairement un despotisme, puisqu'elle établit un pouvoir où tous peuvent décider contre un seul dont l'avis est différent ; la volonté de tous n'y est donc pas exactement celle de tous, ce qui est contradictoire et opposé à la liberté ».

[12] Les chiffres que nous avons donnés sont incertains sur le résultat du face-à-face entre Hollande et Bayrou, mais même si Hollande battait en fait Bayrou, il existerait toujours un vainqueur de Condorcet qui serait alors Hollande lui-même.

[13] On peut imaginer par exemple que les quatre groupes d'électeurs correspondent à des villes parmi lesquelles il faut choisir une capitale (comme dans le dilemme des Alliérins de l'article précédent), chaque électeur voulant que la capitale soit la plus proche possible de sa propre ville.

[14] Noter qu'ici, on montre donc non seulement qu'il existe un vainqueur de Condorcet, mais même qu'il existe un classement global compatible avec les préférences binaires. En fait, un tel classement global existe aussi sous l'hypothèse de préférences en Λ, mais son interprétation n'est pas aussi simple.

[15] À noter que, contrairement à ce qu'on pourrait penser, le fait que toutes les préférences soient en Λ n'entraîne pas que la configuration des préférences est décroisée. Par exemple, si quatre candidats s'appellent de gauche à droite W, X, Y, Z et que deux électeurs ont pour préférences respectives « X>Y>Z>W » et « Y>X>W>Z », la configuration des préférences ne peut pas être décroisée, car la comparaison de X et Y indiquerait que le premier électeur serait plus à gauche que le second tandis que la comparaison de W et Z indiquerait que c'est le second qui serait le plus à gauche.

[16] Confer par exemple le passage de **cette page *Wikipédia*** commençant par « Opinions differ... ».

[17] Cf. M. Balinski, *Le Suffrage universel inachevé* (Belin 2004), pp. 310-311.

[18] Pour être tout-à-fait exact, Condorcet esquissait tout de même le problème dans son essai : il y préconisait apparemment une méthode qu'on appelle aujourd'hui « **méthode Condorcet avec rangement des paires par ordre décroissant** ».

[19] Clin d'œil au célèbre **personnage d'Alan Moore**.

[20] Bien sûr, tout cela est peu réaliste : dans la vraie vie, il faut que les citoyens soient extrêmement mécontents pour oser prendre les armes, et même ainsi les révolutions sont rares ; quant à leur résultat, il dépend de nombreux autres paramètres que le nombre de révolutionnaires ! Mais tout grossier qu'il soit, ce modèle demeure bien adapté au cadre de notre étude.

[21] Vous aurez remarqué que la manipulation des anarchistes, si elle n'a pas permis de porter V au pouvoir, a tout de même changé le vainqueur de Y pour X... Mais on ne peut pas parler de « manipulation » en l'occurrence, vu que ce changement est dû à la modification de votes relatifs à X et à Y.

Affiliation de l'auteur

Rémi Peyre : Institut Élie Cartan (Université de Lorraine) / CNRS

Pour citer cet article : **Rémi Peyre**, **« Et le vainqueur du second tour est... »** — *Images des Mathématiques*, CNRS, 2012. En ligne, URL : http://images.math.cnrs.fr/Et-le-vainqueur-du-second-tour-est.html