LE PARADOXE DE BERTRAND

Tirage au hasard d'une corde dans un cercle.

Jacques HARTHONG, professeur de mathématiques à l'ENSPS de Strasbourg

M. Joseph Bertrand est aujourd'hui peu connu des mathématiciens; il n'y a pas un théorème célèbre associé à son nom, comme le théorème de Cauchy, le théorème de Lebesgue, etc. M. Joseph Bertrand a été un mathématicien, mais ses travaux n'ont pas marqué la postérité de cette science. Pourtant cet homme était très bien connu de tous les mathématiens français de la fin du XIX^e siècle; tous les normaliens ont suivi ses cours; Henri Poincaré, Émile Borel, le citent comme un maître; tout ce qui était édité par l'Académie des sciences portait sa signature. Il a été secrétaire perpétuel de l'Académie des Sciences si longtemps qu'on pourrait appeler le XIX^e siècle mathématique français "siècle de Joseph Bertrand" comme on dit "siècle de Louis XIV".

En 1888 Bertrand publia un livre : *Calcul des probabilités*. Ce livre n'est pas novateur comme l'avait été celui de Laplace sur le même sujet; mais il est très bien écrit et a visiblement servi de référence à une génération. Ce qui est toujours assez surprenant pour un lecteur d'aujourd'hui qui lit ces vieux livres n'est pas seulement la clarté de l'expression et la présence d'*idées* que l'auteur dégage de l'information scientifique, mais aussi (très souvent) la discussion *critique* de cette information scientifique, à laquelle l'auteur convie son lecteur. Le *Calcul des probabilités* de Joseph Bertrand est très représentatif de cet esprit qui s'est perdu : loin de se contenter d'un tableau récapitulatif de tous les résultats techniques connus (qui seraient en quelque sorte garantis par le simple fait qu'ils soient récents), l'auteur discute et critique tout en les exposant les contributions des auteurs antérieurs. Ainsi on trouvera critiquée entre autres un exemple fameux de Laplace sur la probabilité que le soleil se lève demain, sachant qu'il s'est levé tous les jours depuis cinq mille ans, la proposition de Condorcet d'appliquer le Calcul des probabilités aux décisions de justice, un raisonnement célèbre de J. C. Maxwell sur la distribution gaussienne des vitesses des molécules dans un gaz, et (c'est là ce qui nous fournit le sujet de cet article) les fondements des probabilités dites *continues* ou *géométriques*.

J'ajoute encore que certaines des critiques de Bertrand sont elles-mêmes contestables, mais ce n'est pas le sujet de cet article.

À la base de sa critique, il y a chez Bertrand l'idée que les probabilités ne sont pas correctement définies dans le continu. La formule classique

$$\text{probabilité} = \frac{\textit{nombre} \text{ de résultats favorables}}{\textit{nombre} \text{ de résultats possibles}} \tag{1}$$

est claire si les *nombres* qui y figurent sont finis; elle est encore claire comme limite d'une suite convergente lorsque les nombres en question tendent vers l'infini. Mais elle ne l'est plus si dès le départ les nombres sont infinis. Il ne rejette pas en bloc la prise en compte de probabilités géométriques, il met seulement en garde contre les inconsistances qu'on peut y rencontrer.

Parmi les exemples qu'il donne de telles inconsistances, il y a le problème des cordes sur un cercle. Ce problème est entré dans la tradition (on le trouve dans les livres sur les probabilités sous le nom de "paradoxe de Bertrand") depuis que Poincaré l'a repris dans son propre livre *Calcul des probabilités* (1912).

Commençons par exposer le problème puis nous verrons ce qu'en ont dit Joseph Bertrand, puis Henri Poincaré, puis Émile Borel.

Problème : On tire *au hasard* une corde sur un cercle; quelle est la probabilité pour qu'elle soit plus longue que le côté du triangle équilatéral inscrit? (ou, ce qui est équivalent : que la distance de la corde au centre du cercle soit inférieure à $R/2$?)

L'argument (juste) de Joseph Bertrand est que l'ensemble de toutes les cordes possibles forme un continuum; l'ensemble de toutes les cordes favorables, c'est-à-dire de longueur supérieure à $R\sqrt{3}$, est aussi un continuum. On ne peut donc pas appliquer la règle (1). Toutefois, cette critique conduirait à rejeter *tout* problème de probabilités continues, alors que beaucoup d'entre eux (aiguille de Buffon, etc.) ont été résolus avec succès. Bertrand propose donc trois manières, toutes trois conformes aux approches classiques en probabilités géométriques, de résoudre le problème.

1. on tire au hasard deux points sur le cercle (ce qui équivaut à deux nombres compris entre 0 et 2π) : ils définissent une corde, celle qui joint les deux points;

2. on tire au hasard une demi-droite (ce qui équivaut à un nombre θ entre 0 et 2π) et une distance d au centre du cercle (ce qui équivaut à un nombre entre 0 et R); ces deux données définissent univoquement une corde perpendiculaire à la direction donnée et dont la distance au centre est d.

3. on tire au hasard un point à l'intérieur du cercle : ce point définit de manière univoque une corde dont le centre est ce point.

La nouveauté ici, par rapport à l'énoncé initial, est que les ensembles dans lesquels on "choisit au hasard" sont maintenant des ensembles numériques; on choisit à chaque fois un couple de nombres : deux nombres $u_1 \in [0, 2\pi[$ et $u_2 \in [0, 2\pi[$ dans le cas **1**, deux nombres $\theta \in [0, 2\pi[$ et $d \in [0, R]$ dans le cas **2**, et deux nombres x, y (les coordonnées cartésiennes du milieu) dans le cas **3**.

L'ensemble de tous les résultats possibles est donc $[0, 2\pi[\times[0, 2\pi[$ dans le cas **1**, $[0, 2\pi[\times[0, R]$ dans le cas **2**, le disque $x^2 + y^2 < R^2$ dans le cas **3**. Il s'agit toujours d'un continuum, auquel on ne peut appliquer la règle (1). Toutefois, si on ne peut pas définir correctement le *nombre* de cas possibles (qui serait infini), on peut définir l'*aire* des domaines à deux dimensions $[0, 2\pi[\times[0, 2\pi[$ ou $[0, 2\pi[\times[0, R]$ ou encore le disque $x^2 + y^2 < R^2$. On peut donc remplacer la règle (1) par

$$\text{probabilité} = \frac{\textit{aire} \text{ de l'ensemble des résultats favorables}}{\textit{aire} \text{ de l'ensemble des résultats possibles}} \tag{2}$$

Il est facile de caractériser l'ensemble des résultats favorables dans chacun des trois cas : pour que la corde ait une longueur supérieure à $R\sqrt{3}$ il faut et il suffit que
— dans le cas **1** $\frac{2\pi}{3} < |u_2 - u_1| < \frac{4\pi}{3}$;
— dans le cas **2** $d < R/2$;
— dans le cas **3** $x^2 + y^2 < (R/2)^2$.

La réponse est facile à obtenir; impossible de l'exposer plus clairement que Bertrand :

> **cas 1.** On peut dire : si l'une des extrémités de la corde est connue, ce renseignement ne change pas la probabilité; la symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.
>
> L'une des extrémités de la corde étant connue, la direction doit être réglée par le hasard. Si l'on trace les deux côtés du triangle équilatéral ayant pour sommet le point donné, ils forment entre eux et avec la tangente trois angles de $60°$. La corde, pour être plus grande que le côté du triangle équilatéral, doit se trouver dans celui des trois angles qui est compris entre les deux autres. La probabilité pour que le hasard entre trois angles égaux qui peuvent le recevoir le dirige dans celui-là semble, par définition, égale à $\frac{1}{3}$.
>
> **cas 2.** On peut dire aussi : si l'on connaît la direction de la corde, ce renseignement ne change pas la probabilité. La symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.
>
> La direction de la corde étant donnée, elle doit, pour être plus grande que le côté du triangle équilatéral, couper l'un ou l'autre des rayons qui composent le diamètre perpendiculaire, dans la moitié la plus voisine du centre. La probabilité pour qu'il en soit ainsi semble, par définition, égale à $\frac{1}{2}$.
>
> **cas 3.** On peut dire encore : choisir une corde au hasard, c'est en choisir le point milieu. Pour que la corde soit plus grande que le côté du triangle équilatéral, il faut et il suffit que le point milieu soit à une distance du centre plus petite que la moitié du rayon, c'est-à-dire à l'intérieur d'un cercle quatre fois plus petit en surface. Le nombre des points situés dans l'intérieur d'une surface quatre fois moindre est quatre fois moindre. La probabilité pour que la corde dont le milieu est choisi au hasard soit plus grande que le côté du triangle équilatéral semble, par définition, égale à $\frac{1}{4}$
>
> Entre ces trois réponses, quelle est la véritable? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée.

En résumé, la réponse n'est pas la même dans les trois cas : dans le cas 1 elle vaut $\frac{1}{3}$, dans le cas 2 elle vaut $\frac{1}{2}$, dans le cas 3 elle vaut $\frac{1}{4}$.

Bertrand donne cet exemple pour montrer que l'expression “au hasard” n'a pas un sens intrinsèque; il écrit en effet à ce propos :

> L'infini n'est pas un nombre; on ne doit pas, sans explication, l'introduire dans les raisonnements. La précision illusoire des mots pourrait faire naître des contradictions. Choisir *au hasard*, entre un nombre infini de cas possibles, n'est pas une indication suffisante.

Bertrand montre donc que le problème "on tire une corde *au hasard*, quelle est la probabilité pour que ..." est mal posé. Quand on dit "on tire une corde *au hasard*", l'expression *au hasard* devrait signifier que toutes les cordes sont équiprobables, mais on n'a pas une *mesure* sur l'ensemble (continu) de toutes les cordes qui permettrait de connaître le poids de chacune. C'est cette idée de mesure qui va être la clé des deux solutions fournies l'une par Poincaré, l'autre par Borel, au paradoxe. Lorsqu'on réduit (par la modélisation **1**, **2** ou **3**) le problème à des ensembles numériques, ici des domaines du plan, on a une mesure des poids relatifs : si on prend "au hasard" un nombre entre 0 et 2π, la probabilité qu'il soit inférieur à π est $\frac{1}{2}$, etc.

Ainsi la clé du paradoxe est que l'énoncé "on tire une corde *au hasard*, quelle est la probabilité pour que ..." ne fournit pas l'information sur le poids relatif des cordes; on doit ajouter cette information en identifiant chaque corde à un couple de nombres réels, et cette opération comporte un certain arbitraire car il y a plusieurs façons d'identifier chaque corde à un couple de nombres réels.

Dans son propre *Calcul des probabilités* paru une quinzaine d'années plus tard, au chapitre **VII** PROBABILITÉ DU CONTINU, Henri Poincaré expose le problème que nous venons de décrire, avec deux solutions différentes (correspondant aux cas **1** et **2** de Bertrand), sous le titre **Paradoxe de J. Bertrand**. Il conclut :

Pourquoi cette contradiction? Nous avons fait des hypothèses différentes dans les deux cas; nous avons défini la probabilité de deux manières différentes.

D'une manière générale, on demande de définir la probabilité pour qu'un nombre x soit compris entre x_0 et x_1 : en général, nous pouvons dire que nous n'en savons rien du tout.

Cette probabilité doit dépendre de x_0 et x_1 : ce sera donc une fonction telle que $P(x_0, x_1)$.

Ici, Poincaré montre à partir du principe des probabilités totales que cette fonction P sera de la forme

$$P(x_0, x_1) = \int_{x_0}^{x_1} \varphi(x)\, dx \; .$$

puis il continue :

Mais *nous ignorons la nature de* $\varphi(x)$ qui reste arbitraire : il faut nous la *donner* au début du problème par une convention spéciale pour qu'il ait un sens.

De même, la probabilité pour que le point (x, y) soit à l'intérieur d'une aire donnée est

$$\iint \varphi(x, y)\, dx\, dy \; ,$$

l'intégrale double étant étendue à tous les éléments de l'aire; mais nous ne connaissons pas $\varphi(x, y)$.

Le mathématicien n'a plus aucune prise sur le choix de cette hypothèse; mais il doit, une fois qu'elle est choisie, porter son attention à ne pas en faire une autre qui la contredise.

Poincaré explique donc que les deux représentations numériques **1** et **2** de Bertrand pour l'ensemble des cordes correspondent à des choix différents de la fonction φ; la représentation numérique $N°$**3** correspondrait évidemment à un choix encore différent des deux premiers.

Nous y reviendrons plus loin, lorsque nous présenterons le calcul détaillé.

L'analyse de Poincaré repose sur le constat que le choix de φ est une pure convention, même sur des ensembles numériques. La probabilité ne saurait être définie dans l'absolu par la règle (2), il faut instituer une règle plus générale :

$$\text{probabilité} = \frac{\displaystyle\iint_{\{\text{favorables}\}} \varphi(x,y)\,dx\,dy}{\displaystyle\iint_{\{\text{ possibles}\}} \varphi(x,y)\,dx\,dy} \qquad (3)$$

Cet aspect de son approche est très caractéristique : l'insistance sur la convention. On a parlé du *conventionnalisme* de Poincaré, surtout à propos de ses réflexions sur le temps et le principe de relativité de Galilée (cf. *La valeur de la science*, chapitre **II La mesure du temps.**), ou sur l'électromagnétisme (cf. *Théorie mathématique de la lumière*). Il est clair que Poincaré exprime ici la version probabiliste de son conventionnalisme; mais l'idée désignée sous ce terme est trop souvent mal comprise. Les commentateurs sont toujours prompts à classer les idées sous des mots en -isme, mais rarement enclins à les respecter suffisamment pour faire l'effort de les comprendre.

Je demande maintenant au lecteur de ne pas reculer devant cet effort, et de faire pour commencer un peu de mathématique. Nous allons analyser en détail les relations entre les modélisations **1**, **2**, et **3**.

Introduisons les coordonnées suivantes :

- u l'angle polaire de la médiatrice de la corde;
- v le demi-angle d'ouverture de la corde;
- d la distance de la corde au centre du cercle et $t = \dfrac{d}{R}$;
- x, y les coordonnées cartésiennes du milieu de la corde.

On peut dire que dans le modèle 1 on choisit les nombres u et v au hasard, c'est à dire que u est pris au hasard dans l'intervalle $[0, 2\pi]$ et v au hasard dans l'intervalle $[0, \frac{\pi}{2}]$, et on cherche la probabilité pour que v soit dans $[\frac{\pi}{3}, \frac{\pi}{2}]$ (en présentant ce modèle, Bertrand avait dit qu'on choisissait au hasard les deux extrémités u_1 et u_2 de la corde, et non les paramètres u et v introduits ci-dessus, mais nous verrons que cela est équivalent).

Par contre dans le modèle 2 ce sont les nombres u et t qui sont choisis au hasard dans les intervalles $[0, 2\pi]$ et $[0, 1]$, et on cherche la probabilité pour que t soit dans $[0, \frac{1}{2}]$.

Enfin, dans le modèle 3 ce sont les nombres x, y qui sont choisis au hasard dans le domaine $x^2 + y^2 < R^2$, et on cherche la probabilité pour que $x^2 + y^2 < (R/2)^2$.

L'expression *au hasard* signifie que dans les domaines considérés il n'y a pas de région privilégiée et que par conséquent la règle (2) s'applique. La critique de Bertrand était que sur l'ensemble de toutes les cordes, l'expression "au hasard" ne signifie rien, car il n'y a pas de mesure; pour que cette expression prenne un sens, il faut se ramener à des ensembles numériques (ici, des domaines du plan). Nous suivons donc bien ce précepte.

Ces choix *au hasard* se traduisent mathématiquement par le fait que les probabilités sont les rapports des aires, c'est la règle (2) :

cas 1.

$$P_1 = \frac{\int_0^{2\pi} \int_{\frac{\pi}{3}}^{\frac{\pi}{2}} \; dv\, du}{\int_0^{2\pi} \int_0^{\frac{\pi}{2}} \; dv\, du}$$

cas 2.

$$P_2 = \frac{\int_0^{2\pi} \int_0^{\frac{1}{2}} \; dt\, du}{\int_0^{2\pi} \int_0^{1} \; dt\, du}$$

cas 3.

$$P_3 = \frac{\int \int_{x^2+y^2<R^2/4} \; dx\, dy}{\int \int_{x^2+y^2<R^2} \; dx\, dy}$$

Entre les différents paramètres u, v ,t, d, x, y, on a les relations suivantes :

$$\begin{aligned} d &= R\cos v \qquad (\text{et donc} \quad t = \cos v) \\ x &= d\cos u \\ y &= d\sin u \end{aligned}$$

de sorte que si on les exprime tous en fonction de u, t cela donne

$$\begin{aligned} v &= \arccos t \\ x &= R\, t\cos u \\ y &= R\, t\sin u \end{aligned}$$

Si maintenant on rapporte les cas 1 et 3 aux coordonnées u, t, c'est-à-dire si on fait le changement de variable correspondant dans les intégrales, les probabilités deviennent respectivement :

$$P_1 = \frac{\int_{\frac{\pi}{3}}^{\frac{\pi}{2}} \; dv}{\int_0^{\frac{\pi}{2}} \; dv} = \frac{\int_0^{\frac{1}{2}} \frac{1}{\sqrt{1-t^2}} \; dt}{\int_0^{1} \frac{1}{\sqrt{1-t^2}} \; dt} \qquad P_2 = \frac{\int_0^{\frac{1}{2}} \; dt}{\int_0^{1} \; dt} \qquad P_3 = \frac{\int_0^{\frac{1}{2}} t\; dt}{\int_0^{1} t\; dt}$$

(Les intégrales en du se factorisent puis se simplifient, ce qui résulte simplement de la symétrie de rotation).

On voit ainsi que le changement de coordonnées a introduit des densités, $1/\sqrt{1-t^2}$ dans le cas **1** et t dans le cas **3**, qui sont plus grandes pour t proche de 1 que pour t proche de 0. Cela signifie que le choix au hasard du nombre v dans l'intervalle $[0, \frac{\pi}{2}]$ ou le choix au hasard d'un point x, y sur le disque ne correspondent plus à un choix au hasard du nombre t. Si on choisit au hasard v dans l'intervalle $[0, \frac{\pi}{2}]$, de sorte qu'aucune région de cet intervalle ne soit favorisée, alors, pour les valeurs correspondantes de t, on favorise la région proche de 1 par rapport à la région proche de 0.

Cette différence de répartition est due à la non linéarité du changement de variable. Par contre le passage des coordonnées u_1, u_2 (abscisses angulaires des extrémités) aux coordonnées u, v est linéaire : $u_1 = u - v/2$ et $u_2 = u + v/2$; c'est pourquoi dans le cas **1** il est indifférent de choisir u_1, u_2 au hasard ou u, v.

C'est cette différence de répartition qui explique que les probabilités pour que $t < 1/2$ soient plus petites dans les cas 1 et 3 (resp. 1/3 et 1/4) que dans le cas 2 (1/2).

Ainsi *il n'est pas équivalent* de choisir au hasard les deux extrémités de la corde (cas 1), ou de choisir au hasard son milieu (x, y) (cas 3), ou encore de choisir au hasard sa direction et sa distance au centre (cas 2).

L'apparition de ces densités non uniformes montre deux choses :

— a) que, comme Bertrand l'avait souligné, il n'y a pas une uniformité objective sur l'ensemble des cordes, comme il y en a sur l'ensemble des 37 chiffres de la roulette, ou même sur l'ensemble des points d'un intervalle;

— b) que, comme Poincaré l'a souligné, on ne peut pas deviner la densité; si on traite le cas **1** avec les coordonnées u, v on a bien une densité uniforme, et on peut dire qu'on "choisit au hasard" les deux nombres u et v. Mais si on traite le cas **1** avec les coordonnées u, t (ou le cas **2** avec les coordonnées u, v), l'expression de la probabilité sera conforme à la règle (3) qui fait intervenir une densité, et non à la règle (2); cette densité, nous devons la calculer à partir des hypothèses qui caractérisent le modèle **1**, nous ne pouvons pas poser a priori qu'elle est uniforme; ou, si nous le faisons, alors c'est dans l'autre modèle que nous ne pourrons pas poser qu'elle est uniforme.

Pour bien comprendre le point de vue de Poincaré (celui qu'on a appelé le *conventionnalisme*), il ne faut donc pas prendre des phrases telles que :

Mais *nous ignorons la nature de* $\varphi(x)$ qui reste arbitraire : il faut nous la *donner* au début du problème par une convention spéciale pour qu'il ait un sens.

(voir les citations plus haut) et les interpréter hors contexte : elles signifient qu'il ne peut pas y avoir un axiome de la théorie des probabilités qui affirmerait par exemple que la densité φ est uniforme. Il peut arriver qu'elle le soit, mais cela résultera de conditions spécifiques, ce ne sera pas une propriété générale des probabilités. Une phrase telle que celle ci-dessus ne signifie en aucun cas que le mathématicien peut prendre n'importe quelle fonction pour densité sous prétexte que celle-ci serait arbitraire, ni que toutes les densités sont également "vraies" ; elle signifie qu'une théorie *générale* cohérente des probabilités doit prévoir une densité arbitraire pour pouvoir s'adapter à toutes les situations qu'on rencontrera.

Cette phrase ne signifie pas non plus que l'uniformité (l'équiprobabilité) ne joue aucun rôle en Calcul des probabilités et ne serait qu'un cas particulier fortuit; dans l'exemple étudié il est clair que les densités qui apparaissent lors des changements de coordonnées n'auraient pas pu être calculées si on n'avait pas posé les modèles autres que le $N°\mathbf{2}$ en termes d'uniformité. C'est bien l'uniformité selon les variables u et v qui conduit à la densité $1/\sqrt{1-t^2}$ selon les variables u et t. *Ce n'est pas la densité elle-même qui peut être choisie par convention, mais les coordonnées dans lesquelles on décide d'analyser le problème.*

Je me permets d'insister sur ce point. L'interprétation vulgaire du conventionnalisme de Poincaré, c'est-à-dire l'interprétation transmise par ceux qui ne l'ont pas lu et qui se sont contentés d'un vernis scolaire et superficiel, est résumée par le slogan “tout est convention, rien n'est objectif, rien n'est réel”, et apparentée au positivisme. Dans cette interprétation vulgaire, le choix de la densité φ est arbitraire car les trois versions, comme le disait Bertrand, sont également acceptables. Mais cela n'est vrai que pour les problèmes qu'on donne à l'école : le professeur peut alors choisir entre les trois versions; c'est effectivement vrai, mais trivial, et du même niveau que le slogan attribué à Einstein “tout est relatif”. Pour comprendre la véritable *idée* de Poincaré, il faut la mettre dans son contexte, et la rattacher à ses réflexions sur la mesure du temps (cf. *La valeur de la science*) qui sont l'expression la plus célèbre de ce “conventionnalisme”. Poincaré dit

> Le temps doit être défini de telle façon que les équations de la Mécanique soient aussi simples que possible. La mesure du temps est une convention et aucune n'est plus objective qu'une autre.”

Le chapitre “La mesure du temps” dont est extraite cette citation explique comment, de l'astronomie aux vibrations des atomes, les physiciens ont peu à peu fait évoluer la définition du temps pour se rapprocher de l'homogénéité parfaite de son écoulement, dont dérivent la conservation de l'énergie et la notion de repère galiléen. On pourrait prendre à la place du “vrai” temps t n'importe quel paramètre $s = f(t)$ (changement de variable); mais, sauf si f est linéaire, les équations de la Mécanique seraient plus compliquées avec s qu'avec t. On retrouve cela pour les probabilités, comme le montrent les calculs faits plus haut : dans chacun des trois cas étudiés, on a une invariance qui conduit à une équiprobabilité. Si on analyse par exemple le cas **1** dans les coordonnées (u, v), la densité $\varphi(u, v)$ est uniforme; mais si on analyse le même cas **1** dans les coordonnées (u, t), alors $\varphi(u, t) = 1/\sqrt{1-t^2}$. La convention dont parle Poincaré n'est donc pas celle du maître d'école qui peut choisir arbitrairement les énoncés d'exercices, mais celle du choix des coordonnées.

Pour résumer : le conventionnalisme de Poincaré en Calcul des probabilités consiste à dire que le choix des coordonnées dans lesquelles on souhaite décrire un problème donné est libre (il en résultera une densité dépendant de ces coordonnées), et non que les conditions de ce problème sont conventionnelles (ces dernières sont, sauf à l'école, données par la réalité). Exactement comme dans un problème de Mécanique où on est libre de choisir le repère : les *lois* de la Mécanique ne dépendent pas du repère. De même les axiomes d'une théorie des probabilités ne doivent pas dépendre du choix des coordonnées; or, seule la règle (3) est indépendante du choix des coordonnées, et non la règle (2). Ce sont les circonstances pratiques, et non les principes, qui imposent un choix particulier : choisir le repère dans lequel le calcul sera le plus simple possible.

Je regrette de ne pas avoir ici la place pour discuter davantage l'analyse que Poincaré a faite de la mesure du temps en Astronomie et Mécanique, car la similitude est saisissante (cf. *La valeur de la science*).

Pour que les choses soient entièrement clarifiées, je présente encore un problème "concret", dont les conditions ne sont pas conventionnelles mais imposées par la situation. En effet, la raison pour laquelle le point de vue de Poincaré a souvent été mal compris est sans doute que les problèmes scolaires de probabilités sont fabriqués de toutes pièces : si on énonce "on tire une boule au hasard dans une urne" il est sous-entendu par convention que les boules seront équiprobables et on ne se réfère pas à des conditions objectives; d'où le sentiment qu'on peut choisir les hypothèses comme on veut et la confusion de ce sentiment avec le conventionnalisme de Poincaré. Posons le problème des cordes sur un cercle de la manière suivante.

On jette, à partir d'une ouverture pratiquée dans le plafond, des millions de fétus de paille sur le plancher. Afin d'assurer une bonne dispersion des brins, on place un puissant ventilateur près de l'ouverture dans le plafond. Sur le sol, on a tracé un grand cercle à la craie. Pour chaque brin de paille tombé sur le sol, on considère la droite qui le prolonge; dans beaucoup de cas, la droite ne coupera pas le cercle, mais on ne compte pas ces brins. Pour chaque brin dont le prolongement coupe le cercle, on considère la corde que ce prolongement découpe sur le cercle.

Comment se distribuent statistiquement ces cordes?

On a ici affaire à une situation pratique : ce n'est plus *nous* qui décidons d'une manière de choisir les cordes, les cordes se distribuent toutes seules. L'étude des trois cas a montré qu'il y a différentes distributions possibles pour les cordes, mais dans une expérience réelle il y aura forcément une certaine distribution particulière.

Dans cette expérience, c'est le modèle $N°2$ qui est le bon. On peut s'en convaincre sans faire l'expérience, et sans calculer le mouvement des brins. Pour cela, il suffit de remarquer que la distribution des brins sur le plancher ne peut pas dépendre de la présence du cercle : il serait absurde que les brins se distribuent différemment selon qu'on a tracé un cercle ou qu'on n'en a pas tracé (sauf si par exemple le cercle n'était pas matérialisé par de la craie, mais par de petits aimants, qui attireraient les brins de paille supposés aimantés, ou tout autre artifice de ce type). De même, les brins seront distribués à peu près de la même façon dans toutes les régions du plancher. Par conséquent, la distribution des cordes que le prolongement des brins découpe sur un cercle sera la même quelle que soit la région où on trace le cercle. Il est donc impossible que se produisent les situations 1 ou 3, dans lesquelles les cordes sont plus denses près du bord du cercle : si tel était le cas pour un cercle particulier, ce ne pourrait plus l'être pour un autre cercle tracé un mètre plus loin. Il faudrait que les brins "sentent" la présence du cercle et tombent délibérément d'une manière qui favorise le bord du cercle. Seul le modèle $N°2$ correspond à une distribution des droites qui soit indépendante de l'existence du cercle. Ainsi, du fait que les conditions du problème sont déterminées par les circonstances (le ventilateur, les lois du mouvement des brins de paille, etc.), on ne peut "choisir par convention" la loi de probabilité; on peut seulement choisir les coordonnées.

Le paradoxe de Bertrand, l'absence d'une mesure sur l'ensemble des cordes, le caractère arbitraire de la densité φ, tout cela ne signifie pas (comme je l'ai parfois entendu) que le problème ci-dessus aurait plusieurs solutions.

Émile Borel, dans son livre *Le hasard* paru en 1914, reprend le problème de Bertrand et le rediscute. Je rappelle encore ici le sentiment qui pourrait résulter d'une lecture superficielle de Bertrand ou Poincaré, à savoir que les trois réponses au problème seraient toutes également bonnes :

Entre ces trois réponses, quelle est la véritable? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée.

(Joseph Bertrand)

Mais *nous ignorons la nature de* $\varphi(x)$ qui reste arbitraire : il faut nous *la donner* au début du problème par une convention spéciale pour qu'il ait un sens.

(Henri Poincaré)

Voici ce qu'écrit Émile Borel immédiatement après avoir exposé les trois solutions :

Doit-on penser que ces trois solutions sont également bonnes et, par suite, également mauvaises? Nullement, il s'agit simplement de préciser le mode d'après lequel se fera la vérification expérimentale, c'est-à-dire comment on s'y prendra pour tracer une corde *au hasard* dans un cercle : si l'on assujettit cette corde à passer par un point fixe du cercle, ou si l'on fixe son milieu *au hasard*, c'est la première[1] ou la troisième solution qui est la bonne; mais il est aisé de voir que la plupart des procédés naturels que l'on peut imaginer conduisent à la seconde.

Après ces mots, Borel donne un exemple équivalent au problème "concret" précédent (où l'on jette du plafond des fétus de paille). Il conclut :

Comme la solution expérimentale du problème n'est pas contestée on est conduit, dans ce cas pratique, à constater que la seconde des solutions de Bertrand est la seule bonne; il ne nous appartient donc pas, en modifiant arbitrairement une définition, de modifier la solution du problème.

Cette dernière phrase ne contredit pas le conventionnalisme de Poincaré si on a bien compris celui-ci (de toute façon, la version *mal comprise* n'est évidemment pas viable).

Borel conclut que "l'attitude de Bertrand est trop sceptique". Il apparaît en effet que dans son analyse Bertrand conteste la validité générale des probabilités continues; il présente son paradoxe comme un avertissement devant les absurdités qui peuvent surgir à cause du continu. Borel lui reproche de ne pas avoir perçu

[1] Ordre modifié pour le rendre compatible avec notre présentation

l'idée d'invariance, qui existe sur les ensembles numériques, mais pas sur l'ensemble des cordes : pour Borel, c'est l'existence d'une invariance qui fonde l'expression *au hasard*; c'est donc l'absence d'invariance et non le continu qui crée le paradoxe.

Pour bien montrer cela, Borel reprend un autre exemple de Bertrand :

Problème : *Deux points M et M' sont pris au hasard sur la sphère; quelle est la probabilité pour que le plus petit arc de grand cercle MM' soit inférieur à α?*

Bertrand trouve deux solutions incompatibles; nous paraphrasons la présentation de Borel :

Premier raisonnement : La probabilité sera la même quelle que soit la position de M; or, lorsque le point M est fixé, M' doit se trouver sur une calotte sphérique entourant M, correspondant à un demi-angle au centre α. La hauteur de cette calotte est $h = R(1 - \cos\alpha)$, en désignant par R le rayon de la sphère. L'aire de la calotte est proportionnelle à h, donc le rapport de l'aire de la calotte à l'aire de la sphère est

$$\frac{h}{2R} = \frac{1 - \cos\alpha}{2} = \sin^2\frac{\alpha}{2}$$

Telle est la probabilité cherchée; si α est très petit, on peut remplacer le sinus par l'arc et prendre pour valeur approchée $\frac{\alpha^2}{4}$.

Deuxième raisonnement : Lorsqu'on donne deux points M et M' l'arc du grand cercle qui les joint est déterminé (on peut écarter le cas de probabilité nulle où M et M' sont diamétralement opposés); tous les arcs de grand cercle étant analogues sur la sphère, on ne change pas la probabilité en fixant cet arc de grand cercle : or la probabilité pour que deux points M et M' d'un cercle soient tels que l'arc MM' soit inférieur à α est $\frac{\alpha}{\pi}$; ce résultat est très différent du précédent, surtout si α est très petit.

Borel montre que la première solution est conforme aux invariances fondamentales du problème (tous les points de la sphère sont équivalents), mais pas la seconde; en effet, Bertrand dit que, le grand cercle MM' étant fixé, la probabilité que M et M' pris *au hasard* sur *ce* cercle est $\frac{\alpha}{\pi}$. Mais si on garde à l'esprit que la seule invariance *véritable* du problème est l'invariance sphérique, la distribution de probabilité sur un grand cercle donné ne sera pas uniforme le long de ce grand cercle, mais se fera selon une densité φ proportionnelle à la largeur du fuseau formé par deux grands cercles infiniment proches. Comme nous l'avons vu plus haut pour les cordes, cela fait apparaître entre les deux solutions un lien sous la forme d'un changement non linéaire de coordonnées; mais l'argument de Borel est que la première solution est vraiment naturelle, car elle correspond à la “vraie” invariance de la sphère, alors que la seconde ne correspond à rien de naturel :

Doit-on conclure avec Bertrand que le problème proposé ne peut pas être résolu et que la première solution que nous avons donnée est incorrecte? Cette solution est, au contraire, la seule correcte, si l'on admet le postulat relatif à la probabilité élémentaire, c'est-à-dire si on considère toutes les portions égales de la sphère comme équivalentes entre elles, au point de vue de la probabilité qu'a le point M de s'y trouver.

La différence avec les cordes est que sur l'ensemble des cordes, il n'y a pas d'invariance naturelle; il n'y a pas une distribution plus logique qu'une autre. Sur la sphère, si.

On peut dire que Bertrand a certes critiqué des généralisations hâtives et averti son lecteur des absurdités auxquelles on peut être conduit par l'usage irréfléchi de l'expression *au hasard*. Mais il n'a pas cherché à savoir si cette expression avait un sens encore insoupçonné; pour lui, il s'agissait de mots ("la précision illusoire des mots pourrait faire naître des contradictions"). Il réagissait en gardien de la science établie, celle qui a défini ce qui est scientifique et ce qui ne l'est pas, et voulait avertir son lecteur de ce qui n'est pas scientifique.

L'attitude de Borel est toute autre : sa motivation n'était pas de garder ce qui était établi, mais de comprendre ce qu'est le hasard; pour lui l'expression *au hasard* n'était pas un mot illusoire, mais un mystère à percer.

Voyez-vous la différence?

C'est pourquoi Borel a découvert que le hasard est, aussi bien dans le continu que dans le fini, **ce qui choisit entre des possibilités équivalentes**. Et pour cela, il faut qu'il y ait une invariance.

Borel donne dans le même chapitre, après les considérations précédentes, un commentaire sur l'analyse de Poincaré que nous avons vue plus haut (à propos du caractère *arbitraire* de la densité φ), que je conseille très vivement au lecteur; mais je ne la rapporte pas ici, car elle s'écarte de notre sujet (Émile Borel, *Le hasard*, pages 91 – 95)

Reprenons encore le paradoxe des cordes par un autre aspect. La critique de Bertrand était essentiellement fondée sur l'impossibilité de compter l'ensemble de toutes les cordes possibles. Si on prend le terme *ensemble* dans le sens où les mathématiques dites "ensemblistes"[1] l'enseignent, c'est l'ensemble de toutes les cordes distinctes; on peut le mettre en bijection avec l'ensemble des coordonnées u, v ou u, t comme nous l'avons fait, mais (et c'est justement ce qui fait le paradoxe) on ne peut pas définir une mesure sur l'ensemble des cordes par de telles bijections : la mesure dépendrait de la bijection. La critique de Bertrand ne peut évidemment pas concerner des ensembles finis, où tout se passe sans paradoxe ni ambiguité. Que se passerait-il alors si on discrétisait l'ensemble des cordes? Pour discrétiser un ensemble continu on représente par un seul élément toute une famille continue d'éléments proches de celui qu'on a choisi.

On voit déjà poindre ce qui va se passer : le mot "proche" invoque une métrique, c'est-à-dire une mesure : dira-t-on que la corde C est proche de la corde C_0 si ses extrémités sur le cercle diffèrent de moins de ε de celles de C_0? Ou si sa direction et sa distance au centre du cercle diffèrent de moins de ε de celles de C_0?

[1] Et je rappelle que Borel a participé à la construction de la théorie des ensembles.

Ou si son centre est à une distance inférieure à ε de celui de C_0? On retrouvera respectivement les trois modèles.

Il est immédiat que les trois modèles ont l'invariance par rotation; la discrétisation par rapport à la coordonnée u est donc commune aux trois modèles; seule la partie radiale de la discrétisation les distinguera. Sur les figures 1, 2, et 3, on a représenté cette partie radiale de l'ensemble discrétisé; il faut imaginer que l'ensemble discrétisé complet est formé par la superposition de 360 figures radiales, tournées chacune de un degré par rapport à la précédente.

Les trois ensembles de la figure ont le même nombre d'éléments; mais ces éléments ne sont pas distribués de la même façon le long du rayon : ils correspondent aux densités $1/\sqrt{1-t^2}$ (fig 1), 1 (fig 2), et t (fig 3). Le nombre d'éléments situés à moins de $R/2$ du centre est donc plus faible sur les figures 1 et 3 que sur la figure 2.

Il est bien clair que sur chacun de ces ensembles finis, où chaque corde est équiprobable, il n'y a pas d'ambiguité concernant le calcul de la probabilité; la règle (1) s'applique. L'ambiguité qui avait fait le paradoxe se reporte sur la discrétisation : il y en a une différente pour chaque modèle. On comprend en quoi le continu masque cela : sur le continuum de cordes, on ne voit pas la densité des cordes; sur les figures la densité est le nombre de cordes par millimètre; mais dans le continu elles sont toutes collées ensembles et on ne peut pas voir si elles sont rares ou denses.

Je finirai par une remarque, qui éclaircira encore le conventionnalisme de Poincaré : il y a une analogie très forte entre l'équiprobabilité des choix du hasard et le principe d'inertie en Mécanique :

— Une transformation linéaire des coordonnées conserve l'équiprobabilité (dans le modèle **1** la densité est uniforme aussi bien dans les coordonnées u, v que dans les coordonnées u_1, u_2) tout comme une transformation galiléenne conserve les mouvements rectilignes uniformes.

— Une transformation non linéaire des coordonnées fait apparaître une densité φ, tout comme un repère non galiléen fait apparaître des mouvements non inertiels.

— Un problème de probabilités peut être analysé avec n'importe quelles coordonnées (chacun des modèles **1**, **2**, **3** peut indifféremment être analysé dans les coordonnées u, v ou u, t ou encore x, y), tout comme un problème de Mécanique peut être analysé dans n'importe quel repère non nécessairement galiléen.

— Enfin, et surtout : quelqu'un qui observe la distribution des cordes sur une figure du type 1, 2, ou 3 peut *en déduire* la manière dont "le hasard" a choisi les cordes (s'il a choisi au hasard les coordonnées u, v dans $[0, 2\pi[\times[0, \frac{\pi}{2}]$, ou s'il a choisi au hasard les coordonnées u, t dans $[0, 2\pi[\times[0, R]$, ou encore s'il a choisi au hasard les coordonnées x, y dans le disque), tout comme quelqu'un qui observe un champ de pesanteur constant dans une capsule spatiale située loin de toute masse peut en déduire que celle-ci est uniformément accélérée.

Le conventionnalisme de Poincaré consiste à dire qu'il n'y a pas de coordonnées privilégiées dans l'absolu, et qu'il faut choisir celles qui conduisent aux calculs les plus simples possibles.

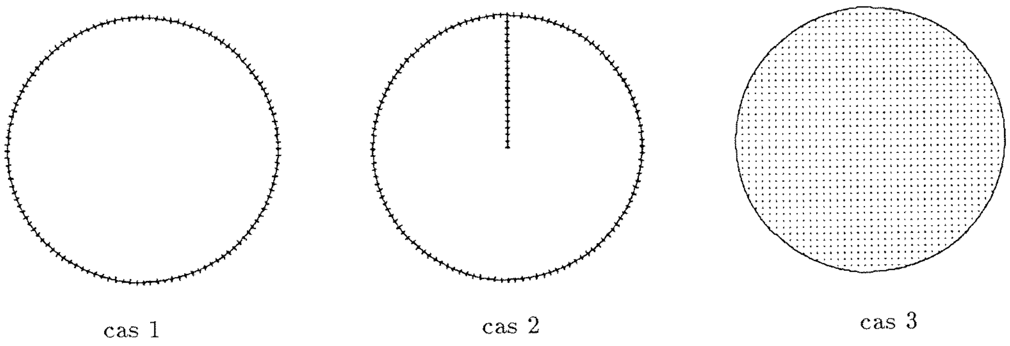

figure 0

On a représenté la discrétisation du problème dans les trois cas. Afin de ne pas surcharger la figure, on n'a représenté la discrétisation que radialement pour les cas 1 et 2. Dans le cas 3, le quadrillage détruit bien sûr la symétrie sphérique, mais celle-ci est conservée statistiquement.

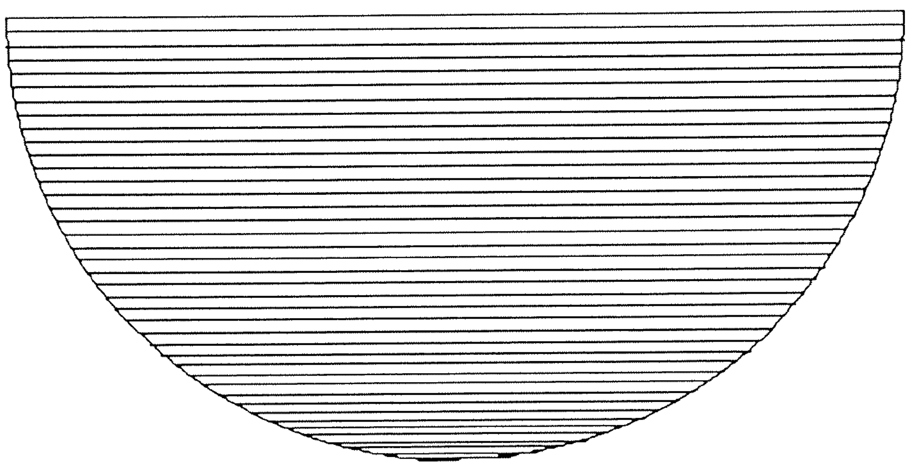

figure 1

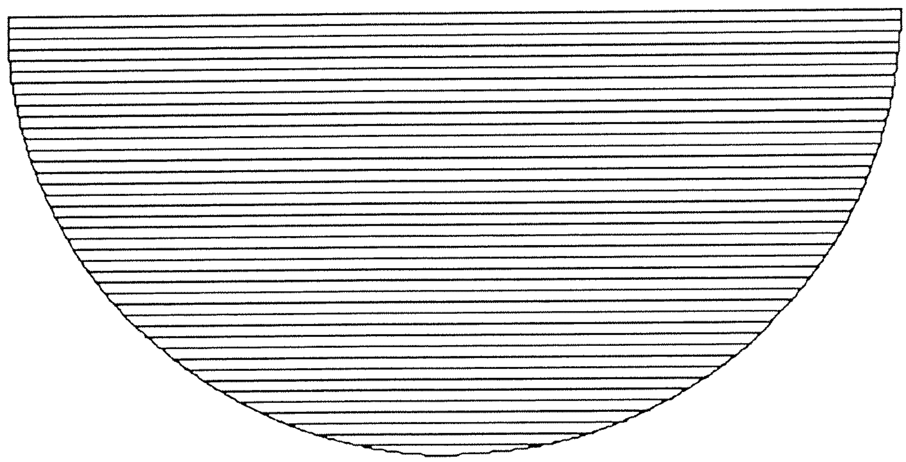

figure 2

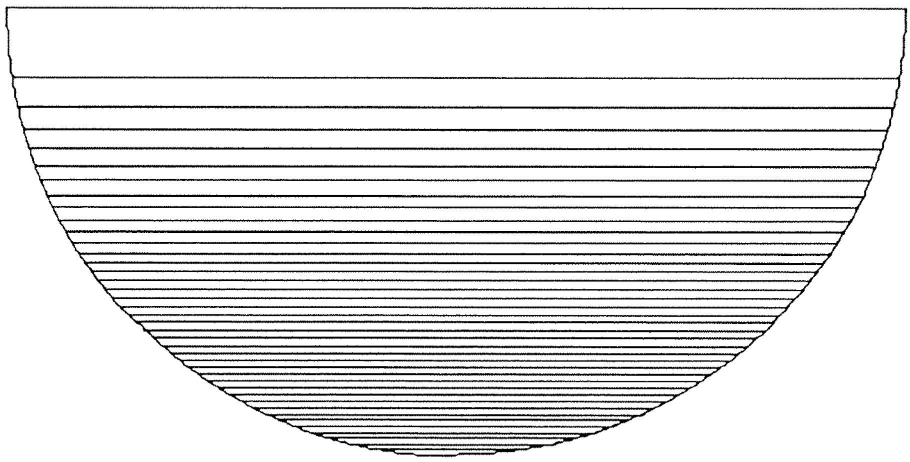

figure 3