

# I. LE HASARD.

## I. 1. Le langage du Calcul des probabilités.

“La théorie des hasards consiste à réduire tous les événements du même genre à un certain nombre de cas également possibles et à déterminer le nombre de cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles est la mesure de cette probabilité, qui n'est ainsi qu'une fraction dont le numérateur est le nombre de cas favorables, et dont le dénominateur est le nombre de tous les cas possibles.”

Pierre-Simon Laplace

*Théorie analytique des probabilités* (1819).

Il n'existe pas de meilleure définition de la probabilité que celle-ci, vieille de plusieurs siècles, et c'est d'elle que nous partirons.

Par exemple si on tire “au hasard” un nombre entre 1 et 36, la probabilité d'obtenir 11 est  $\frac{1}{36}$ ; si la probabilité d'obtenir 11 était  $\frac{1}{32}$ , qui est supérieur à  $\frac{1}{36}$ , cela signifierait que le nombre 11 est privilégié par la Fortune ou que la roulette est truquée. Il y aurait une cause agissant secrètement et le hasard pur ne pourrait éventuellement être retrouvé qu'au-delà de cette cause. On ne pourrait pas considérer que le tirage est soumis au seul hasard.

On appelle *épreuve* un tel résultat possible. Par principe, toutes les épreuves sont supposées équiprobables.

N.B. Lorsqu'on applique le Calcul des probabilités à la réalité il est bien rare que l'on ait le loisir de s'assurer rigoureusement et à l'avance que cette hypothèse d'équiprobabilité est bien vérifiée, car d'innombrables *causes* peuvent intervenir à notre insu pour fausser les chances. Le modèle mathématique postulant l'équiprobabilité peut être valable lorsque les différentes causes sont suffisamment nombreuses et de forces égales pour qu'aucune ne domine (par exemple dans le cas des accidents de la route), ou bien que des symétries fondamentales de l'espace-temps assurent a priori qu'il ne peut y avoir de domaine privilégié (par exemple dans le cas des phénomènes physiques). Dans le jeu de la roulette, c'est la conception même de la roulette qui garantit l'équiprobabilité avec une grande précision. Toutefois, nous verrons (§2) que l'hypothèse d'équiprobabilité ne signifie rien en soi: encore faut-il savoir quelles sont les épreuves équiprobables.

Un résultat attendu peut être tout un ensemble d'épreuves. Pour donner un exemple, revenons au tirage d'un nombre entre 1 et 36: on peut se demander quelle est la probabilité d'obtenir un nombre premier. La réponse est  $\frac{11}{36}$ . Le seul moyen de trouver cette probabilité est de faire la liste exhaustive des nombres premiers compris entre 1 et 36: 2, 3, 5, 7, 11, 13,

17, 19, 23, 29, 31. On constate qu'ils sont au nombre de onze, soit un peu moins que le tiers de 36. Le résultat attendu n'était pas une épreuve isolée, mais tout un ensemble d'épreuves réunies par une propriété commune.

On appelle *événement* un tel ensemble d'épreuves. L'équiprobabilité des épreuves a pour conséquence immédiate que la probabilité d'un événement est le rapport :

$$\frac{\text{nombre d'épreuves appartenant à l'événement}}{\text{nombre de toutes les épreuves possibles}} \quad (I.1)$$

Dans l'exemple des nombres premiers, le seul moyen de connaître la probabilité était de compter exhaustivement. Mais dans beaucoup de problèmes comportant des régularités, des symétries, ou des invariances, une méthode mathématique appropriée permet de calculer, sans avoir à compter. C'est l'ensemble de ces méthodes qui constitue le *Calcul de Probabilités*.

L'ensemble de toutes les épreuves possibles est évidemment spécifique du problème étudié. On l'appelle *espace des épreuves* (en anglais *sample space*) et on le note très souvent  $\Omega$ . Dans le langage mathématique on écrit

$$\begin{aligned} \Omega = \{ & 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, \\ & 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, \\ & 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 \} \end{aligned}$$

Le sous-ensemble des nombres premiers est, dans le même langage mathématique

$$A = \{ 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31 \}$$

L'événement  $A$  est ainsi un sous-ensemble de  $\Omega$ , qui est généralement défini par une propriété commune à tous ses éléments, à savoir ici, "être un nombre premier". Dans le langage du Calcul des probabilités, on désigne souvent un événement par une phrase imagée qui exprime cette propriété commune à tous ses éléments. On dira que  $A$  est l'événement "on a tiré un nombre premier". Ainsi on retrouve le sens du mot événement dans le langage courant. Dans le langage du Calcul des probabilités, on désigne indifféremment un événement par une phrase imagée exprimant une certaine propriété, ou comme un sous-ensemble de  $\Omega$ . Le sous-ensemble en question est alors l'ensemble des épreuves qui possèdent ladite propriété. La phrase imagée permet de se faire comprendre facilement, elle exprime aussi le sens du problème; la traduction mathématique sous forme d'ensemble sert par contre à calculer. Il faut s'exercer à passer de l'un à l'autre. Nous verrons beaucoup d'exemples qui permettront de s'y entraîner.

Si  $A$  est un événement ( $A \subset \Omega$ ) on note  $\#A$  le nombre d'éléments (le cardinal) de l'ensemble  $A$ , c'est-à-dire le nombre d'épreuves qui constituent l'événement  $A$ . La probabilité d'un événement  $A$  est ainsi :

$$\mathcal{P}(A) = \frac{\#A}{\#\Omega} \quad (I.2)$$

qui est une traduction immédiate de (I.1.)

La résolution de n'importe quel problème concret comporte traditionnellement deux étapes :

- a) la modélisation (nom moderne de ce qu'on appelait autrefois "mise en équations")
- b) la résolution analytique ou le traitement numérique de ces équations.

En Calcul des probabilités, la phase a) de modélisation consiste à trouver l'espace des épreuves  $\Omega$  correspondant au problème; la phase b) consiste à calculer les probabilités à l'aide de I.2.

La théorie mathématique des probabilités développe une *algèbre*, c'est-à-dire des règles de calcul, qui s'appliquent en général aux probabilités des événements dans des espaces d'épreuves non nécessairement finis. Pour des espaces d'épreuves *finis* (le seul cas qui nous intéresse) cette théorie est élémentaire, mais suffit à traiter tous les problèmes. La suite de l'ouvrage traitera de problèmes classiques qu'il faut connaître, de méthodes de calcul approché, de lois asymptotiques, d'applications diverses.

Ainsi le langage du Calcul des probabilités est extrêmement simple. Pour le moment nous en avons présenté les bases, qui se réduisent au vocabulaire suivant.

- **épreuve** : un des résultats possibles, que le hasard peut choisir, sans en favoriser aucun.
- **espace des épreuves** : l'ensemble de toutes les épreuves.
- **événement** : un ensemble d'épreuves, généralement défini par une propriété commune à ses éléments.
- **cardinal** : nombre d'éléments d'un ensemble fini.
- **probabilité** : la probabilité d'un événement  $A$  est le rapport du cardinal de  $A$  au cardinal de  $\Omega$ . C'est donc un nombre toujours compris entre 0 et 1.

Peu à peu ce vocabulaire de base sera enrichi par de nouvelles définitions.

Le mot *hasard* ne fait pas partie du vocabulaire technique du Calcul des probabilités, car ce dernier ne se préoccupe que des règles et des méthodes de calcul. Une discussion sur la nature du hasard est cependant essentielle;

sans cela, le Calcul des probabilités n'est qu'un formalisme vide, qu'on saura utiliser pour résoudre des problèmes scolaires énoncés dans le même formalisme, mais qu'on ne saura plus utiliser devant des problèmes concrets *dépourvus* de formalisation préalable. Nous avons dit que la *modélisation* d'un problème de probabilités consistait à trouver l'espace des épreuves adéquat. Mais "espace des épreuves" sous-entend équiprobabilité, et il n'est pas évident de savoir ce qui est équiprobable. C'est cela qui soulève toute une discussion sur la nature du hasard, à laquelle le reste de ce chapitre est consacré.

### I. 2. Étude de deux exemples.

L'exemple extrêmement simple du tirage au hasard d'un nombre entre 1 et 36 ne pose aucun problème de modélisation : le choix de l'ensemble  $\Omega$  adéquat est immédiat. Il ne laisse rien soupçonner de la profondeur du concept. C'est pourquoi nous étudions dans cette section deux problèmes non triviaux (quoique simples), afin d'illustrer dès maintenant le véritable sens des concepts déjà introduits.

#### Premier exemple : distribution au hasard de trois boules dans deux boîtes.

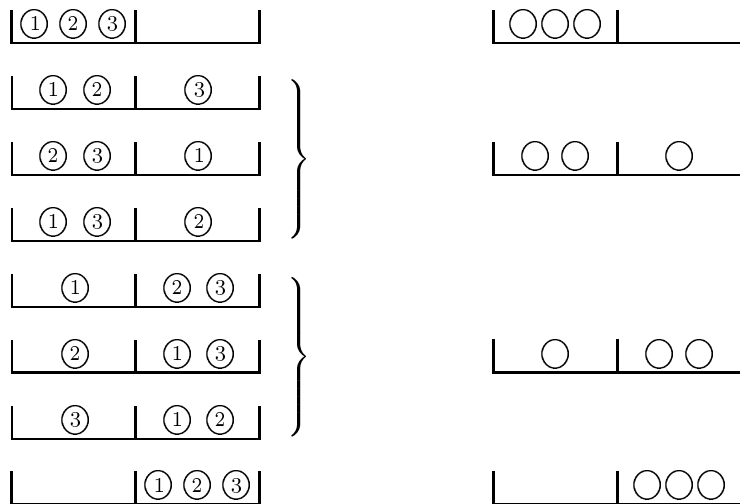


figure 1

On voit sur la figure 1 : dans la colonne de gauche, toutes les répartitions possibles de trois boules numérotées 1, 2, 3 dans les deux boîtes ; il y en a huit en tout. Si on supprime les marques sur les boules, on ne peut plus distinguer les répartitions (1,2 | 3), (2,3 | 1), et (1,3 | 2), ni (1 | 2,3),

(2 | 1, 3), et (3 | 1, 2). C'est pourquoi, sur la colonne de droite, il n'y a plus que quatre répartitions.

Question : pour des boules sans marques, l'espace des épreuves est-il formé des huit répartitions de la colonne de gauche, ou des quatre de la colonne de droite ?

Il se trouve que la réponse dépend de la nature des boules : s'il s'agit de boules macroscopiques obéissant à la Mécanique classique, c'est la première réponse qui convient (et peu importe qu'elles soient numérotées ou non, car le problème n'est pas qu'elles soient subjectivement numérotées, mais qu'elles soient objectivement numérotées) ; s'il s'agit de particules quantiques obéissant à la statistique de Bose-Einstein, c'est la deuxième réponse qui convient. L'équiprobabilité est bien dans les deux cas l'expression du *pur hasard*, mais il se trouve que le pur hasard ne se situe pas au même niveau de causalité selon que les particules sont classiques ou quantiques.

Dans la plupart des ouvrages sur le Calcul des probabilités, on ne donne pas de la probabilité  $\mathcal{P}(A)$  une définition directement dérivée de l'hypothèse d'équiprobabilité comme nous le faisons ici, mais une définition axiomatique du type suivant : "étant donné un ensemble fini  $\Omega$ , appelé espace des épreuves, on appelle *probabilité* sur  $\Omega$  une application  $\mathcal{P}$  de l'ensemble des parties de  $\Omega$  sur l'intervalle  $[0, 1[$  qui vérifie la propriété d'additivité  $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B)$  pour tout couple de parties  $A$  et  $B$  disjointes, ainsi que la propriété  $\mathcal{P}(\Omega) = 1$ ". Il est bien clair que la probabilité telle que nous l'avons définie dans (I.2.) satisfait à cette définition axiomatique. Mais ce que la définition axiomatique recouvre est plus général et n'exige pas nécessairement que toutes les épreuves soient équiprobables ; on peut donner à chaque épreuve un *poids*, c'est-à-dire un atome de probabilité. Si les poids sont différents il n'y a pas équiprobabilité et (I.2.) doit être remplacé par

$$\mathcal{P}(A) = \frac{\text{somme des poids des épreuves appartenant à } A}{\text{somme des poids de toutes les épreuves}} \quad (II.2 \text{ bis.})$$

Ceci vérifie aussi la définition axiomatique et inversement, toute application de l'ensemble des parties de  $\Omega$  sur l'intervalle  $[0, +\infty[$  qui vérifie la propriété d'additivité sera de la forme (II.2 bis.), avec le poids  $\mathcal{P}(\{\omega\})$  pour chaque épreuve  $\omega$ .

Les auteurs qui procèdent ainsi ont pour cela une excellente raison : exposer le Calcul des probabilités pour des épreuves non nécessairement équiprobables ne coûte pas une seule phrase de plus que pour des épreuves équiprobables, car tout se déduit de l'additivité (ce qu'on déduit de I.2. passe toujours par l'intermédiaire logique de l'additivité). Pourquoi donc se priver de quelque chose d'absolument gratuit ?

En fait l'enjeu est le *sens* du Calcul des probabilités, ou autrement dit, la signification du mot *hasard*.

La caractéristique du pur hasard est l'équiprobabilité. Nous reviendrons encore sur ce point au paragraphe suivant. Si des épreuves ne sont pas équiprobables, c'est que certaines sont favorisées et donc qu'une cause, qui n'est pas le hasard, agit secrètement. Il ne s'agit pas là d'un choix philosophique : si une étude statistique établit que le cancer du pancréas est plus fréquent dans la haute vallée de l'Isère que partout ailleurs, on en déduit qu'une cause, agissant dans cette haute vallée et non ailleurs (ou agissant plus fortement dans cette haute vallée qu'ailleurs) est responsable de cette surfréquence. Même si après de

longues recherches cette cause reste introuvable, il n'y aura pas de querelle philosophique sur la question de savoir si la Fortune est ou n'est pas aveugle; on ne parlera de hasard que si la fréquence est distribuée uniformément, car c'est là le sens même du mot hasard.

Il peut être commode de disposer de modèles théoriques où l'équiprobabilité n'est pas postulée a priori. Par exemple dans un problème où on jette  $r$  boules dans  $k$  cases identiques, mais où on ne s'intéresse qu'à l'alternative: "case 1" ou "autre case", il est préférable de modéliser la situation en ne parlant que de deux possibilités seulement, mais de sorte que la probabilité de la première soit  $1/k$  et la probabilité de la deuxième  $1 - 1/k$ . Il s'agit alors d'un modèle phénoménologique et on retrouve le vrai hasard en creusant davantage. Il faut cependant admettre que dans ce cas le gain en simplicité qui résulte du modèle phénoménologique (comparé au modèle avec équiprobabilité) est médiocre. Une expérience plus poussée du Calcul des probabilités montrerait que le gain est médiocre dans tous les cas. Nous verrons au chapitre **IV** la *formule des probabilités conditionnelles* qui établit le lien entre de tels modèles phénoménologiques et le modèle équiprobable sous-jacent.

En revanche le fait de poser les problèmes de probabilité en remontant systématiquement au *pur hasard*, sans s'arrêter à des représentations phénoménologiques, est foncièrement instructif.

C'est justement ce que montre l'exemple ci-dessus de trois boules à disposer dans deux cases. On peut pour les deux situations discutées considérer un espace des épreuves à quatre éléments, les quatre éléments étant :

$$\begin{aligned} \omega_1 &: \begin{cases} \text{trois boules dans la case 1} \\ \text{zéro boule dans la case 2} \end{cases} \\ \omega_2 &: \begin{cases} \text{deux boules dans la case 1} \\ \text{une boule dans la case 2} \end{cases} \\ \omega_3 &: \begin{cases} \text{une boule dans la case 1} \\ \text{deux boules dans la case 2} \end{cases} \\ \omega_4 &: \begin{cases} \text{zéro boule dans la case 1} \\ \text{trois boules dans la case 2} \end{cases} \end{aligned}$$

Dans le cas de la statistique de Bose-Einstein, les boules sont par exemple des photons et les cases des états quantiques; les quatre épreuves  $\omega_1, \omega_2, \omega_3, \omega_4$  sont équiprobables; dans le cas des boules macroscopiques, par contre (et peu importe qu'elles soient numérotées ou non, elles sont toujours *objectivement* discernables, du fait qu'elles sont et restent situées dans l'espace-temps), les quatre épreuves ne sont pas équiprobables; leurs probabilités respectives sont  $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$ . On peut alors calculer tout ce qu'on veut sur les probabilités de tous les événements concernant ces épreuves, et on aboutira aux mêmes résultats dans le modèle à quatre éléments non équiprobables que dans le modèle à huit éléments équiprobables. Mais bien entendu, il aura fallu auparavant trouver les valeurs  $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$ . Et pour cela il n'y avait aucun autre moyen que de sortir du modèle à quatre éléments, venir en secret dans le modèle à huit éléments équiprobables pour y déterminer les valeurs des poids, puis retourner au modèle à quatre éléments en faisant semblant de les avoir devinées grâce au génie.

Si on analyse bien cet exemple, on comprend que le *pur hasard* est une expression des symétries fondamentales: si les huit épreuves du modèle à huit éléments sont effectivement équiprobables c'est parce que les trois boules peuvent être isolées dans l'espace. Quoiqu'il arrive à l'une des trois boules, cela n'affecte pas l'état des deux autres. Lorsque la Fortune dispose la deuxième dans l'une des deux cases, sa cécité consiste à ne pas percevoir l'état d'occupation des deux cases par la première boule, mais seulement les positions spatio-temporelles relatives de la deuxième boule et des

deux cases. En revanche, lorsque la même Fortune dispose le second photon dans l'un des deux états, sa cécité consiste à ne pas percevoir les positions spatio-temporelles relatives des photons (ces positions n'ont d'ailleurs aucun sens objectif) et des deux états quantiques, mais seulement les états d'occupation. Ainsi dans le premier cas aucune des boules ne favorise une case, tandis que dans le second cas aucun des photons ne favorise un état d'occupation. Dans le premier cas le pur hasard est le reflet d'une invariance spatio-temporelle (les deux cases sont équivalentes car ce sont des lieux symétriques de l'espace, et chacune des trois boules agit indépendamment des autres, c'est-à-dire que sa trajectoire ne dépend pas du fait que les autres boules ont été lancées avant ou après). Dans le second cas le pur hasard est le reflet d'une indiscernabilité objective des trois photons; seuls peuvent être objectivement distingués les quatre états d'occupation.

Or aucun problème de probabilité ne peut être résolu sans recourir directement au postulat que la Fortune est aveugle. Si une possibilité est favorisée par rapport à une autre (dans le sens qu'elle est plus probable) on ne peut calculer a priori sa probabilité qu'en recherchant le hasard pur à un niveau supérieur. Mais on ne peut évidemment appliquer le postulat que si on sait quelle symétrie intervient dans le problème. Cela est tellement vrai qu'on peut le vérifier par exemple dans le livre de William FELLER *An Introduction to Probability Theory and its Applications* (John Wiley ed.) Je choisis ce livre parce que c'est, sur le sujet, le livre de référence: il est pratiquement impossible d'enseigner mieux que FELLER les questions traitées dans son livre; c'est pourquoi certaines parties de ce cours sont simplement reprises d'après FELLER (voir par exemple le chapitre *marches aléatoires*). Bien qu'ennemi juré du formalisme mathématique, FELLER ne renonce pas à la commodité de l'approche axiomatique qui, comme nous l'avons dit plus haut, ne coûte aucun effort supplémentaire, et laisse la porte ouverte aux espaces d'épreuves non équiprobables. De ce fait il est conduit à donner des exemples ou des exercices où interviennent des espaces d'épreuves non équiprobables. Mais bien sûr dans chacun de ces exemples ou exercices, les poids qui sont "parachutés" au lecteur ont été déterminés par l'auteur à partir d'un modèle à épreuves équiprobables.

Ainsi, exercice 4. page 24:

"A coin is tossed until for the first time the same result appears twice in succession. To every possible outcome requiring  $n$  tosses attribute probability  $1/2^{n-1}$ . Describe the sample space. Find the probability of the following events: (a) the experiment ends before the sixth toss, (b) an *even* number of tosses is required."

Voir aussi les deux exercices qui suivent celui-ci.

Le choix du poids  $1/2^{n-1}$  provient d'une équiprobabilité à un niveau supérieur (nous reviendrons là-dessus au chapitre *probabilités conditionnelles*). On trouvera encore un tel exemple pages 47 – 48 (chap *II* §7: Examples for waiting times). Fréquemment dans son ouvrage, FELLER exprime la condition "assuming perfect randomness, ...". Cette condition signifie toujours que le problème possède un modèle avec équiprobabilité, et dans lequel cette équiprobabilité résulte d'une symétrie évidente (le tout étant de la trouver). On appellera *probabilités a priori* les probabilités ainsi obtenues par considération d'une symétrie. En l'absence de symétrie connue on peut aussi mesurer a posteriori les probabilités par la statistique; on ignore alors où est intervenu le pur hasard, et on parlera de *probabilités empiriques*. On peut toujours construire a posteriori un modèle phénoménologique reprenant ces probabilités empiriques. Mais j'ai estimé préférable de recourir à la formule des probabilités conditionnelles (chap. **IV**).

**Deuxième exemple: tirage au hasard d'une corde dans un cercle.**

Cet exemple est célèbre: il est tiré du *Calcul des probabilités* de Joseph Bertrand, un ouvrage paru en 1888. Sa célébrité provient de ce qu'il a été repris et commenté par Henri Poincaré, puis encore par Émile Borel. Il examine trois manières différentes de tirer au hasard une corde dans un cercle de rayon  $R$ :

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

2. on tire au hasard une direction (ce qui équivaut à un nombre  $\theta$  entre 0 et  $2\pi$ ) 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 à la distance  $d$  du centre;

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

À première vue le problème relève du Continu: il y a un continuum de cordes, et les nombres qui les déterminent sont des nombres réels. Cela ne pose aucune difficulté particulière, mais il est de toute façon aisé de discrétiser le problème: on découpe les intervalles dans lesquels les nombres doivent être choisis au hasard en un grand nombre de parties égales. Cela est montré sur la figure 2 pour les trois cas.

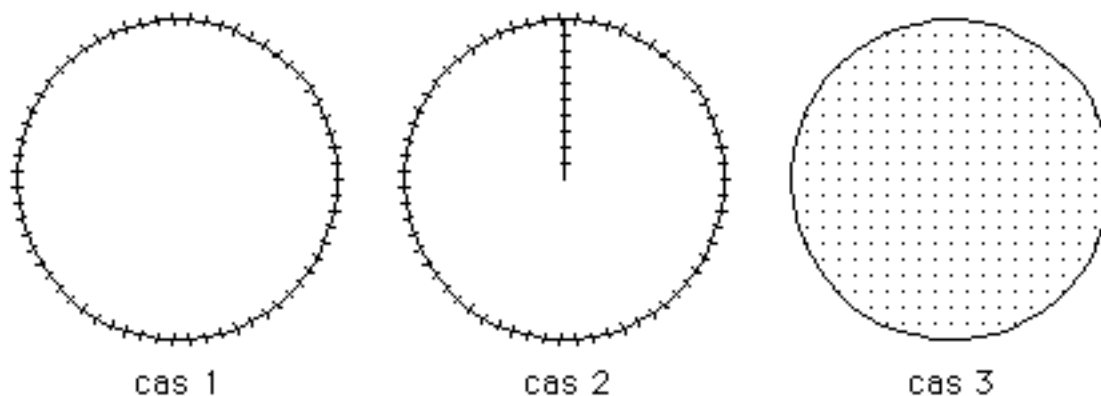


figure 2

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.

**Problème:** quelle est la probabilité pour que la distance de la corde au

centre du cercle soit inférieur à  $R/2$  ou ce qui est équivalent, quelle est la probabilité pour que la longueur de la corde soit supérieure à  $R\sqrt{3}$ ?

Les réponses dépendent du modèle: dans le cas 1 elle vaut  $\frac{1}{3}$ , dans le cas 2 elle vaut  $\frac{1}{2}$ , dans le cas 3 elle vaut  $\frac{1}{4}$ . La différence provient du fait que l'espace  $\Omega$  de toutes les cordes possibles n'est pas le même dans les trois cas. Les figures 3.1, 3.2, et 3.3 montrent quel est l'espace  $\Omega$  respectivement dans les cas 1, 2 et 3 (on n'a représenté que la partie radiale). Dans son livre, Joseph Bertrand présentait cet exemple dans un but de critique, comme un problème mal posé. Il avait raison, puisque l'énoncé ci-dessus ne permet pas à lui seul de déterminer lequel des trois modèles est le bon.

Une analyse détaillée du phénomène permet de comprendre la différence entre les trois figures. 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  $t = \frac{d}{R}$

$x, y$  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  $[0, \frac{\pi}{2}]$  (en présentant ce modèle, nous avons 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). L'expression *au hasard* signifie que dans ces intervalles il n'y a pas de région privilégiée. Si on discrétise le problème comme dans la figure 2, cela conduit à diviser ces intervalles en parties égales.

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  $[0, 1]$ .

Enfin, dans le modèle 3 ce sont les nombres  $x, y$  qui sont choisis au hasard dans le domaine  $x^2 + y^2 < R^2$ .

Ces choix *au hasard* se traduisent mathématiquement par le fait que les probabilités sont les rapports des aires :

**cas 1.**

$$P_1 = \frac{\int_0^{2\pi} \int_0^{\frac{\pi}{2}} dv du}{\int_0^{2\pi} \int_0^{\frac{\pi}{2}} dv du}$$

**cas 2.**

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

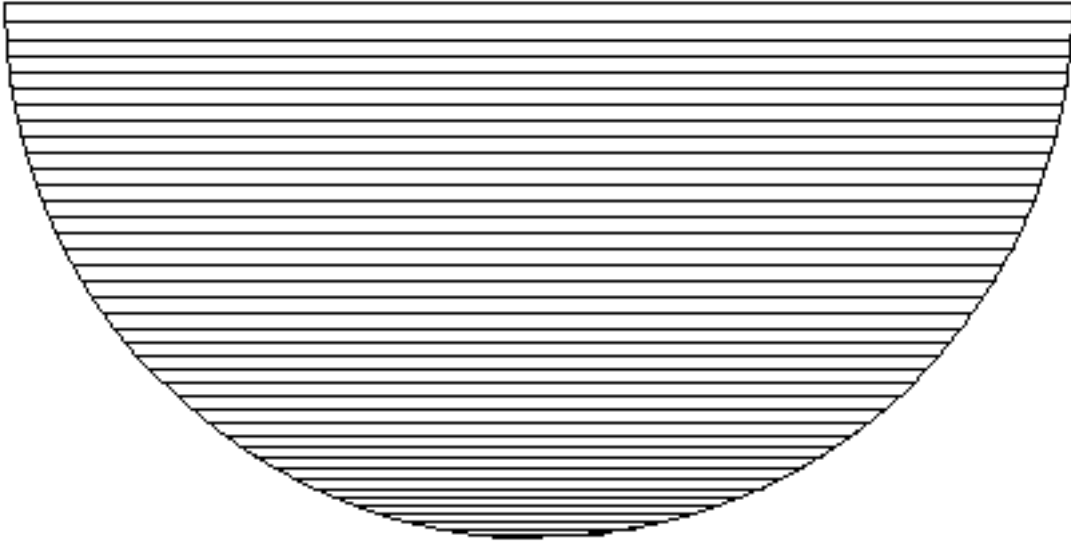


figure 3.1.

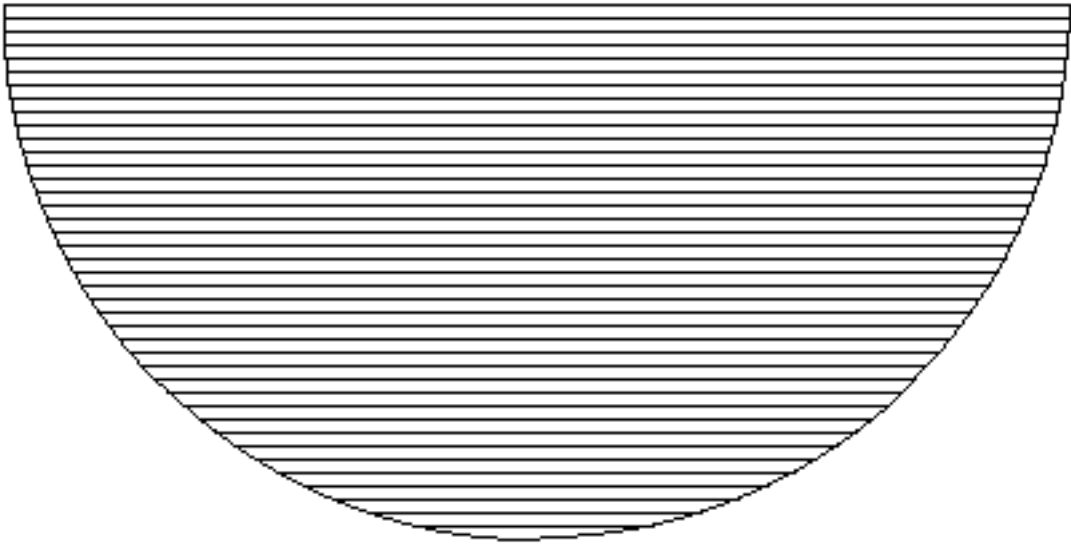


figure 3.2.

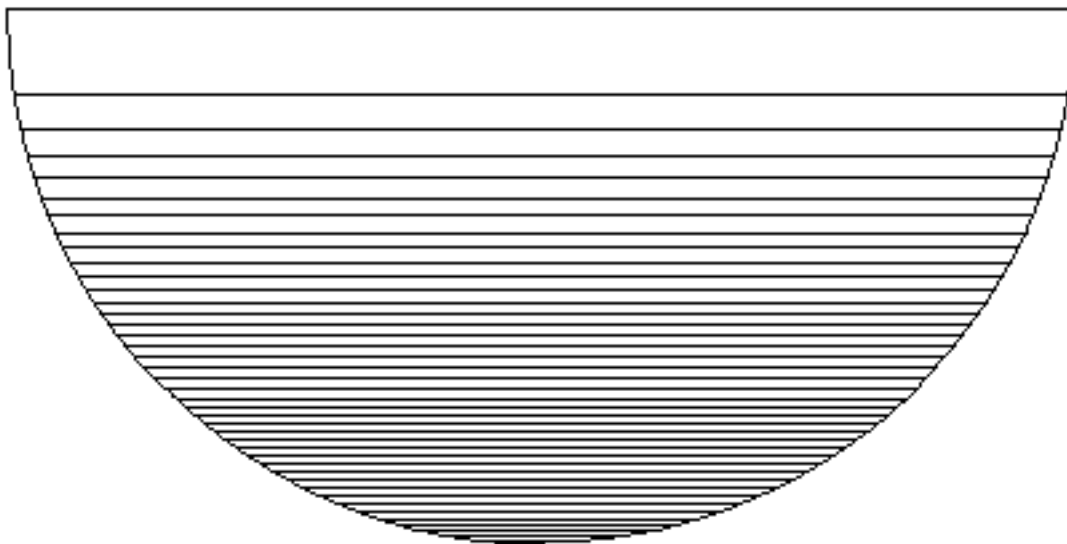


figure 3.3

cas 3.

$$P_3 = \frac{\iint_{x^2+y^2 < R^2/4} dx dy}{\iint_{x^2+y^2 < R^2} dx dy}$$

Entre ces différents paramètres on a les relations suivantes :

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

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 :

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

(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 deux, 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 cercle 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. Elle apparaît clairement sur les figures 3.1 et 3.3 : les cordes proches du bord (ce qui correspond à  $t$  proche de 1) sont favorisées par rapport à la figure 3.2. 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 ce 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 ( $1/2$ ). Joseph Bertrand attribue l'ambiguïté au continu ; il est vrai que lorsqu'on discrétise, comme sur les figures 3, on voit bien la différence, qui ne se verrait pas dans le continu. Cependant Émile Borel<sup>(1)</sup> a critiqué cette critique de Bertrand en montrant que l'ambiguïté ne vient pas du continu, mais du fait que le niveau où intervient le hasard n'est pas précisé.

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). Mais supposons qu'on fasse l'expérience concrète que voici : 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, 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 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.

---

<sup>(1)</sup> Émile Borel *Le hasard* (Librairie Félix Alcan, 1914), page 84

Où intervient alors le hasard? 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  $N^{\circ}2$  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 (sauf bien sûr si on s'écarte trop du trou dans le plafond). 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  $N^{\circ}2$  correspond à une distribution des droites qui soit indépendante de l'existence du cercle.

### **I. 3. La signification de l'équiprobabilité.**

Jusqu'ici nous avons souvent évoqué le "pur hasard", qui ferait des "choix équiprobables". Ces considérations sur le *pur hasard* ne sont pas nécessaires pour un exposé axiomatique et déductif, qui suppose un modèle mathématique (c'est-à-dire un espace des épreuves  $\Omega$ ) déjà donné, et récuse toute question sur l'origine ou sur le sens de ce modèle. Elles le deviennent cependant si on veut maîtriser l'art de la modélisation et être capable devant une situation concrète de construire soi-même le modèle adéquat.

L'affirmation que *le pur hasard* fait des choix équiprobables n'est pas un postulat métaphysique. C'est un *principe*, qui est au Calcul des probabilités ce que le principe de relativité de Galilée est à la Mécanique. Toutefois il peut laisser le sentiment qu'il s'agit de métaphysique à cause de l'invocation du "pur hasard". Ce n'est pas le mot hasard qui insulte ainsi le positivisme, mais plutôt l'épithète qui l'accompagne et qui évoque "les pures idées", "la Raison pure", etc. C'est pourquoi l'explication qui va suivre est nécessaire.

Afin que la discussion soit bien concrète, voyons comment les choses se passent dans l'exemple simple des cordes sur un cercle. Nous avons vu

à ce propos qu'il y avait différentes manières de choisir une corde sur un cercle. Imaginez maintenant un physicien qui observe pendant longtemps l'apparition aléatoire de cordes sur un cercle, mais qui ignore "comment le hasard a choisi les cordes" (vous pouvez imaginer que c'est un programme écrit par vous-même qui fait apparaître, au rythme de une par seconde, des cordes sur l'écran de l'ordinateur; pour le choix de chaque corde le programme fait appel deux fois – puisqu'il faut deux coordonnées – à une fonction **random**; le physicien qui observe le résultat ignore comment sont écrits les programmes). Il accumule les observations et, sachant qu'il y a la symétrie de rotation, il ne note la position des cordes que si, à  $\varepsilon$  près, elles sont parallèles à une direction fixée; il obtient ainsi, disons, la figure 3.1. Ce physicien peut *en déduire* que les cordes ont été distribuées en laissant le hasard choisir les deux extrémités, indépendamment l'une de l'autre, sur le cercle. Dans le graphique qu'il obtient (semblable, donc, à la figure 3.1), les cordes ne sont pas distribuées uniformément le long des ordonnées, c'est-à-dire le long du rayon. "Le hasard pur" n'est donc pas intervenu à ce niveau (celui de la distribution le long du rayon, qui est aussi le choix de la coordonnée  $t$ ) mais au niveau du choix des extrémités. (si les cordes sont tracées par votre programme, le physicien peut ainsi savoir que la fonction **random** choisit les deux extrémités, comme dans le cas 1, et non les coordonnées  $x, y$  du centre ou les coordonnées  $t, u$ ). Le physicien peut donc, par l'observation, départager les différentes manières d'agir du hasard. Cela montre qu'il s'agit d'une information pratique et observable, et non d'une hypothèse métaphysique.

Lorsqu'on dit "le hasard pur est intervenu à ce niveau" ou "le hasard pur choisit les deux extrémités", cette manière imagée de s'exprimer peut suggérer qu'il y aurait une divinité douée de volonté et veillant, par une vigilance de tous les instants, à ne favoriser aucun de ses choix par rapport aux autres, et que par conséquent le problème posé ci-dessus (trouver le niveau où le hasard fait son choix), consisterait à pénétrer les desseins de cette divinité. Il ne s'agit que d'une façon de parler, comme par exemple lorsqu'on dit "la Nature obéit à des lois".

On ne postule aucune divinité, et on peut au contraire postuler que le hasard provient de causes toutes rationnelles, qu'il n'est que l'effet d'un enchevêtrement d'innombrables influences qui se contrarient mutuellement, créant ainsi l'illusion appelée hasard. Mais si ces causes, rationnelles ou non, favorisent certains résultats au détriment d'autres, on dira que ce favoritisme a une explication causale et non qu'il est un effet du hasard.

Pour illustrer cela, revenons encore à l'exemple des cordes sur un cercle. Dans ce problème il n'y avait pas de phénomène physique; il s'agissait de géométrie abstraite et "les choix du hasard" portaient sur des nombres.

Ainsi “le hasard” devait choisir, soit deux nombres  $\alpha, \beta$  dans l’intervalle  $[0, 2\pi[$ , soit deux nombres  $u, t$  dans (respectivement) les intervalles  $[0, 2\pi[$  et  $[0, 1]$ , soit un point dans le disque. Pour tout calculer, nous sommes partis du seul postulat que quand “le pur hasard” choisit un nombre dans un intervalle, il ne favorise aucune région de l’intervalle: par exemple les nombres de  $[0, 1]$  n’ont pas plus de chances d’être choisis s’ils sont proches de 0, que s’ils sont proches de 1. Ce qu’on exprime mathématiquement en disant que la probabilité pour que le nombre choisi dans  $[0, 1]$  se trouve “par hasard” entre  $a$  et  $b$  est égale à  $b - a$ . Il n’y a rien de plus dans cette expression mathématique que l’intuition première, que tout le monde comprend a priori. De même pour un point “pris au hasard” dans le disque: la probabilité pour que “par hasard” il se trouve dans la région  $A$  du disque est le rapport

$$\frac{\text{aire de } A}{\text{aire du disque}}$$

Or l’étude détaillée des différentes formes de choix a montré que si par exemple les deux extrémités  $\alpha, \beta$  de la corde sont choisies dans l’intervalle  $[0, 2\pi[$ , sans favoriser aucune région de cet intervalle, alors *ipso facto* la coordonnée  $t$  (distance au centre rapportée au rayon) sera choisie – dans l’intervalle  $[0, 1]$  – de manière que les nombres proches de 1 soient favorisés par rapport à ceux qui sont proches de 0. Mathématiquement, la probabilité pour que  $t$  se trouve “par hasard” entre  $a$  et  $b$  n’est plus  $b - a$ , mais

$$\int_a^b \frac{dt}{\sqrt{1-t^2}}$$

Le physicien que nous évoquions plus haut, qui observe ce favoritisme, ne dira pas (*personne* ne le dira) “le hasard préfère les cordes proches du bord que les cordes proches du centre”. Il dira “il y a un choix aléatoire, *biaisé* par un phénomène déterministe”.

Attention ! Il ne s’agit pas d’une querelle de mots. Si on change les mots, on ne supprime pas l’insatisfaction du physicien, qui veut savoir pourquoi les valeurs de  $t$  ont plus de chances d’être proches de 1 que de 0. On pourrait *convenir* – dans le sens où Henri Poincaré<sup>(2)</sup> disait que la mesure du temps est une convention et qu’aucune n’est plus objective qu’une autre – que le hasard choisit  $t$  de façon non équiprobable, avec la densité  $1/\sqrt{1-t^2}$ . Mais cela ne nous épargnerait pas les questions, cela ne ferait que nous rendre l’ensemble des calculs effectués pour les trois cas plus compliqués, car nous devrions traîner partout la répercussion de cette densité, de même que si on voulait absolument mesurer le temps selon une échelle différente, et non

---

(2) Henri Poincaré *La valeur de la science*, chapitre II “La mesure du temps”.

linéaire par rapport à l'usuelle, on devrait traîner la répercussion de ce changement de variable dans toutes les équations de la Physique. Poincaré a certes écrit que la mesure du temps est une convention et qu'aucune n'est plus objective qu'une autre, mais la phrase qui est écrite immédiatement avant dans son texte est "le temps doit être défini de telle façon que les équations de la Mécanique soient aussi simples que possible".

C'est de là que dérive la notion de repère galiléen, et c'est pourquoi "le niveau où les choix du hasard sont équiprobables" joue en Calcul des probabilités un rôle analogue à celui du repère galiléen en Mécanique. Lorsque le physicien observe les cordes, distribuées le long du rayon selon la loi  $1/\sqrt{1-t^2}$ , il peut *en déduire* que le niveau où les choix étaient équiprobables était celui du choix des extrémités de la corde, de même qu'un physicien enfermé dans une capsule loin de toute source de gravitation, qui observe que les billes qu'il lâche dans sa cabine suivent un mouvement uniformément accéléré, peut en déduire que sa capsule subit de l'extérieur un mouvement uniformément accéléré en sens opposé à celui des billes.

Afin de ne pas trop surcharger ce chapitre, la discussion sur le "conventionnalisme" de Poincaré, ses propres commentaires sur le paradoxe de Bertrand, et la parenté avec le principe de relativité de Galilée ne sont ici que résumés succinctement. Le sujet est davantage développé dans un article, "*Le paradoxe de Bertrand*", qui peut être trouvé à l'adresse U.R.L. :

[http : //moire4.u-strasbg.fr/hist/bertrand.htm](http://moire4.u-strasbg.fr/hist/bertrand.htm)

Tout cela est bien beau, direz-vous, mais ne nous dit pas comment on va trouver le niveau où agit le hasard pur, c'est-à-dire le niveau où on a affaire à des choix équiprobables. En outre, il pourrait y avoir des niveaux différents, mais équivalents, et on ne pourrait pas savoir dans lequel des deux le hasard agirait "vraiment".

Il se trouve qu'il n'existe pas de méthode pour cela. Quant à la question de trancher entre deux niveaux équivalents, elle par contre est métaphysique; elle est exactement aussi métaphysique que la question de savoir quel repère galiléen est "vraiment" immobile. Dans le cas des cordes sur un cercle, la recherche du niveau où agit le pur hasard a été facile, car nous connaissions déjà trois différentes manières de distribuer les cordes et il a suffi d'en reconnaître une parmi les trois. Dans les problèmes d'urnes, de boules, de marches aléatoires, etc. (autrement dit dans les problèmes scolaires de Calcul des probabilités) cette recherche du "niveau où agit le hasard pur" relève de la *modélisation*; le "niveau où agit le hasard pur" est très souvent suggéré, ou même tout simplement donné, par l'énoncé du problème.

La modélisation a toujours fait peur aux étudiants car elle demande de

l'imagination, de l'astuce, et sa réussite n'est pas prévisible ou chiffrable; tandis que si la modélisation est déjà faite, ou suggérée de façon évidente, il n'y a plus qu'à calculer en appliquant quelques théorèmes ou formules vues dans le cours. Tout cela est classique.

Mais il y a plus grave. Après tout, lorsqu'un étudiant n'arrive pas à trouver le modèle correct, c'est parce qu'il n'est pas encore habitué aux poncifs des problèmes scolaires, et dans ce cas l'enseignant connaît la bonne réponse, qu'il avait lui-même apprise quand il était étudiant, par mimétisme plus que par imagination. Par contre il y a des problèmes posés par la Physique pour lesquels *personne* ne sait au départ à quel niveau agit le hasard pur, et le trouver est une affaire de génie (ce qui justifie tout-à-fait, d'ailleurs, le malaise que les étudiants éprouvent face à la modélisation). L'exemple historique le plus spectaculaire est évidemment la découverte de la Mécanique quantique : à partir de faits expérimentaux (rayonnement du corps noir, diffraction des électrons par des cristaux, etc.) il a fallu trouver par la recherche le niveau auquel le hasard pur agissait ; on ne l'a toujours pas vraiment trouvé. Il serait donc absolument naïf de vouloir disposer d'une méthode générale et systématique pour le trouver.

Il semble ressortir de cette discussion que l'équiprobabilité des choix du hasard, tout comme pour la mesure du temps ou les repères galiléens, est en quelque sorte *la définition même* du hasard. On pourrait penser que le hasard peut se définir autrement, indépendamment de la manière dont s'opèrent ses choix, à partir de considérations plus fondamentales, de telle sorte que l'équiprobabilité serait une propriété qui se déduirait de la définition, et non un postulat a priori. Il se trouve que pour le temps ou les repères galiléens, la Physique ne fournit pas de connaissance plus profonde que ce conventionnalisme (c'est le nom donné à la position de Poincaré citée précédemment), du moins jusqu'à nos jours. En revanche, l'Histoire de la Physique contient beaucoup d'exemples où un progrès dans la connaissance apporte une explication plus approfondie d'un phénomène dont la description avait été jusque là purement conventionnaliste. Ainsi nous comprenons aujourd'hui *pourquoi* le mouvement apparent des planètes dans le firmament (c'est-à-dire par rapport à la *sphère des fixes*) est comme il est ; par exemple nous comprenons pourquoi une planète avance (dans le même sens que les étoiles), puis ralentit, puis revient en arrière, selon une trajectoire apparemment erratique : l'explication en est que nous la voyons en perspective à partir de la Terre qui est elle-même en mouvement.

Cette explication est l'*idée* du système de Copernic. Mais dans le système de Ptolémée, où la Terre était supposée immobile au centre, il fallait supposer plusieurs sphères imbriquées les unes dans les autres et tournant autour d'axes différents, la plus éloignée étant la sphère des fixes qui portait les

étoiles, les sphères intérieures portant chacune une planète. Dans ce système de Ptolémée le mouvement erratique des planètes était certes expliqué par les mouvements relatifs des différentes sphères, mais il n’y avait pas *une* combinaison unique de sphères compatible avec le mouvement observé. De sorte que dans ce système la question de savoir laquelle de ces combinaisons était “réellement” en fonction dans le ciel était une question métaphysique ; on ne pouvait répondre que par le conventionnalisme, en disant que parmi toutes les combinaisons de sphères produisant les mouvements planétaires apparents observés, il fallait choisir celle pour laquelle les calculs sont les plus simples possibles. Dans l’astronomie moderne, au contraire, il n’y a pas de choix conventionnel : le système solaire est tel qu’il est, nous ne savons pas pourquoi il est précisément comme cela plutôt qu’autrement, mais nous ne pouvons pas décider de décrire autrement la trajectoire de Mars dans le seul but de simplifier les calculs. Si nous n’avons plus cette liberté, c’est parce que nous en savons trop : les astronomes du Moyen-Age considéraient que seul le mouvement apparent des planètes dans le ciel (donc sa projection radiale sur la voûte céleste) était une connaissance objective. Aujourd’hui leur mouvement dans les *trois* dimensions est soumis à la connaissance objective, ce qui détruit toute possibilité de conventionnalisme à ce niveau. L’Histoire de la Physique regorge d’autres exemples semblables.

C’est pourquoi la question posée plus haut à propos du hasard mérite d’être approfondie. Le hasard peut-il être conçu comme objectif, de sorte que l’équiprobabilité de ses choix soit une nécessité, ou au contraire, à l’instar du temps et de l’espace, est-il – du moins jusqu’à nouvel ordre – impensable autrement que comme convention ?

Il se trouve que, selon les présupposés, l’une ou l’autre réponse est possible. C’est ce que nous allons examiner dans les deux sections suivantes.

#### **I. 4. Hasard et déterminisme mathématiques.**

Tous les événements, ceux mêmes qui par leur petitesse semblent ne pas tenir aux grandes lois de la nature, en sont une suite aussi nécessaire que les révolutions du soleil. Dans l’ignorance des liens qui les unissent au système entier de l’univers, on les a fait dépendre des causes finales, ou du hasard, suivant qu’ils arrivaient et se succédaient avec régularité, ou sans ordre apparent ; mais ces causes imaginaires ont été successivement reculées avec les bornes de nos connaissances, et disparaissent entièrement devant la saine philosophie, qui ne voit en elles que l’expression de l’ignorance où nous sommes des véritables causes.

Pierre-Simon Laplace

*Théorie analytique des probabilités* (1819).

Ce passage célèbre est le manifeste du déterminisme : rien n’est aléatoire, tout est rigoureusement déterminé, mais souvent le développement déterministe d’un processus est tellement chaotique que nous ne pouvons rien

calculer du processus exact, que seules des probabilités peuvent être calculées. La présente section et la suivante sont consacrés à cette explication du hasard. L'affirmation de Laplace n'a jamais été démentie depuis, puisque là où elle pourrait éventuellement l'être nous sommes ignorants : c'est ce qu'on appelle un postulat métaphysique (il ne peut être tranché par l'expérience). Avec la Mécanique quantique nous *ne savons pas* d'où provient le hasard qu'elle postule. Mais par exemple le hasard dans le jeu de la roulette obéit entièrement à la conception de Laplace. La question de savoir si le hasard postulé par la Mécanique quantique peut s'expliquer à la manière de Laplace est ouverte, mais il faut dire que tant que la valeur d'une telle explication ne peut pas être tranchée par l'expérience, elle reste spéculative. Peut-être une telle explication, même sans être tranchée par l'expérience, pourrait-elle du moins réduire le mystère de ce hasard dont nul ne comprend la cause ; mais pourquoi un déterminisme serait-il moins mystérieux que le hasard ?

Aujourd'hui on appelle *chaos déterministe* le mécanisme par lequel, comme l'entendait Laplace, le déterminisme se transforme en hasard. Mais il faut aussitôt ajouter que le déterminisme aussi peut être l'effet du hasard, par l'intermédiaire de la loi des grands nombres (nous reviendrons là-dessus aux chapitres suivants, notamment à la fin du chapitre **II**, ainsi qu'aux chapitres **VI** et **VII**). De sorte que, du déterminisme ou du hasard, chacun peut être la cause première de l'autre.

Comme nous l'avons toujours fait jusqu'ici (et, rassurez-vous, nous continuerons) nous allons rendre la discussion concrète par l'étude de quelques exemples simples.

Voici un premier exemple extrêmement simple de chaos déterministe : on prend un nombre irrationnel, disons  $e \simeq 2.718\,281\,828\dots$  pour fixer les idées. Puis on considère la suite de nombres  $u_n$  = la partie décimale de  $n \cdot e$  (c'est-à-dire qu'on ne retient du nombre  $n \cdot e$  que les chiffres après la virgule). Ainsi les nombres  $u_n$  sont tous compris entre 0 et 1. Pour donner une idée, voici les cinq premiers termes de cette suite :

$$u_1 = 0.718\,281\,828\dots$$

$$u_2 = 0.436\,563\,656\dots$$

$$u_3 = 0.154\,845\,485\dots$$

$$u_4 = 0.873\,127\,313\dots$$

$$u_5 = 0.591\,409\,142\dots$$

Si on calcule un très grand nombre de termes de cette suite, disons un milliard, on obtiendra donc un gros échantillon de nombres compris entre 0 et 1.

Il se trouve que cette suite de nombres a apparemment les mêmes propriétés qu’une suite tirée au hasard ; d’après le préjugé que nous avons sur le hasard, à savoir qu’il ne favorise pas une région particulière de l’intervalle,  $10^9$  nombres “pris au hasard” se répartissent uniformément dans l’intervalle : dans n’importe quel sous-intervalle  $[a, b] \subset [0, 1[$ , il y en aura environ  $(b - a) \cdot 10^9$ . Mais il en va de même pour la suite déterministe  $u_n$  : ces nombres se répartissent uniformément dans l’intervalle. Si, en évitant soigneusement de révéler comment la suite  $u_n$  a été engendrée, on en confiait le résultat à un statisticien pour qu’il le soumette aux tests usuels de la statistique, il serait impossible à celui-ci de détecter un écart par rapport aux lois du hasard ; le seul écart, infime, qui pourrait être observé, diminuerait en augmentant la taille de l’échantillon (nous verrons au chapitre **XI** comment on peut effectuer de tels tests). Dans l’exemple des cordes sur un cercle, si au lieu de confier le choix des coordonnées au hasard, on avait tiré les nombres selon la suite  $u_n$  (ou  $2\pi u_n$  pour les nombres devant être pris entre 0 et  $2\pi$ ), tout se serait passé exactement de la même façon. Cette définition mathématique du hasard vient d’Émile Borel<sup>(3)</sup>

Plusieurs fois nous avons évoqué l’emploi d’une fonction **random** pour simuler le hasard, ce qui fait évidemment plus sérieux que d’évoquer la Fortune. Mais les fonctions **random** du commerce fonctionnent exactement sur le même principe que la suite  $u_n$  ; tout au plus elles font appel à des algorithmes plus sophistiqués afin de mieux brouiller les pistes<sup>(4)</sup>. On va voir en effet que la suite  $u_n$  ne brouille pas suffisamment bien les pistes.

Supposons que, au lieu d’avoir pris le nombre  $e \simeq 2.718281828$ , nous ayons pris le nombre 0.5. La suite des  $v_n =$  partie décimale de  $n \cdot 0.5$  ne respecte pas la règle de ne favoriser aucune région, puisqu’elle ne donne que 0 ou 0.5, et défavorise donc les nombres situés dans  $[0.1, 0.4]$  ou dans  $[0.6, 0.9]$ . Dans ce cas on ne peut plus parler de hasard. La différence entre les deux situations ne provient pas du déterminisme, qui serait absent dans un cas et présent dans l’autre (et si même il y avait indéterminisme, on ne serait pas plus avancé, puisqu’on aurait juste remplacé le mot hasard par le mot indéterminisme). Il y a déterminisme dans les deux cas, mais dans le premier la règle de l’uniformité est respectée, tandis que dans le second elle ne l’est pas. Toutefois, si le statisticien auquel on a confié la suite  $u_n$  pour expertise, considère la suite  $w_n = u_{n+1} - u_n$ , il constatera que les

<sup>(3)</sup> Émile Borel *Les probabilités dénombrables et leurs applications arithmétiques* (Rendiconti del Circolo matematico di Palermo, vol. **27** (1909, pp. 247 – 270).

<sup>(4)</sup> On peut cependant considérer que plus l’algorithme est tortueux et cryptique, plus le hasard qui en résulte est de bonne qualité. Ce point de vue est à la base de la théorie des suites aléatoires. Voir par exemple le livre de Donald Knuth *The Art of Computer Programming* Vol. 2 : Seminumerical Algorithms (Addison-Wesley, 1969) ou celui de Jean-Paul Delahaye *Information, complexité, et hasard* (Éd. Hermès, Paris, 1994).

nombres  $w_n$  ont le même défaut que les  $v_n$ , car ils sont tous égaux, soit à  $e - 2 \simeq 0.718\,281\,828$ , soit à  $e - 3 \simeq -0.281\,718\,171$ .

On ne peut donc pas vraiment admettre que la suite  $u_n$  soit au hasard car “il fallait avoir fait exprès” de choisir les  $u_n$  pour que leurs différences soient toujours les mêmes. On a donc en tête le préjugé que le hasard ne peut pas avoir choisi des nombres que l’application d’un algorithme aurait également pu engendrer (nous avons déjà le préjugé que le hasard ne favorise pas certains nombres au détriment d’autres; voilà donc un autre préjugé). On pourrait alors donner du hasard la définition suivante: un ensemble de  $N$  nombres réels pris dans l’intervalle  $[0, 1]$  sont distribués au hasard si leur liste ne peut être engendrée par aucun algorithme. L’ennui, c’est que (du moins pour un nombre fini  $N$  de nombres) il y a toujours un algorithme, ne serait-ce que la donnée directe de ces  $N$  nombres. On peut corriger cette incohérence en précisant davantage: on dira que  $N$  nombres réels de l’intervalle  $[0, 1]$  sont distribués au hasard si leur liste ne peut être engendrée par aucun algorithme *plus court que la simple liste de ces  $N$  nombres*. Pour bien comprendre cela, imaginons qu’on veuille écrire un programme qui écrit les  $N$  nombres dans un fichier de données. Pour  $N = 10^9$  et huit octets par nombre (il s’agit de flottants), ce fichier est très gros: 8 gigaoctets. Mais le programme qui engendre la suite  $u_n$  est très court, car l’algorithme est très simple:

```
var i : integer ;
    N : longinteger ;

begin
  N := 1e+9 ;
  for i := 1 to N do
    begin
      writeln(i * exp(1) - trunc(i * exp(1))) ;
    end
  end.
```

Ce programme, même en code exécutable, va occuper au plus 1000 octets de mémoire. Ce qui permet cette énorme économie par rapport à la *liste* des nombres, est la boucle qui n’est écrite qu’une fois, bien que dans l’*exécution* du programme elle soit répétée  $10^9$  fois. Si les  $10^9$  nombres n’étaient liés entre eux par aucune récurrence, on ne pourrait pas utiliser de boucle; le programme devrait comporter autant d’entrées ou d’initialisations que de nombres, et serait donc aussi long que la liste des nombres.

Autrement dit la suite  $u_n$  peut être écrite par un programme court car il suffit de connaître  $u_1 = e$  pour déduire par récurrence tous les autres. Si les  $10^9$  nombres n’étaient liés entre eux par aucune relation, il ne suffirait

pas d'en connaître un ou même quelques uns pour en déduire les autres, de sorte qu'on ne pourrait pas trouver de programme *plus court* que la liste directe.

Ce qui a permis à notre statisticien de découvrir la tricherie, est la possibilité de faire apparaître le déterminisme de la suite par les soustractions. Mais si *le plus court* algorithme qui engendre la suite est horriblement long, celui-ci échappera forcément aux investigations et aux tests. On peut donc mesurer le degré de hasard ou – inversement – le degré de déterminisme d'une suite de nombres à la longueur en *Ko* du plus court programme qui engendre la suite. Ainsi un degré élevé de hasard signifie uniquement que la suite est engendrée par un algorithme suffisamment tortueux pour pouvoir échapper à la surveillance des inspecteurs. Cette idée découle naturellement de celle de Borel mentionnée plus haut, mais elle n'a été érigée en système que plus tard, et est le résultat d'une lente maturation qui s'étale de 1900 à 1960 environ. Le lecteur intéressé pourra lire le livre de Knuth ou celui de Delahaye (déjà cités en note 4) pour en savoir plus.

La conclusion ultime de toute cette lente maturation d'idées sur ce que seraient “vraiment” des nombres pris au hasard dans un intervalle, est qu'il n'y a pas de “vrai” hasard, il n'y a que des algorithmes déterministes, mais suffisamment cryptés et tortueux pour que *en pratique* aucun expert ne puisse découvrir le déterminisme sous-jacent. C'est ainsi, en tous cas, que fonctionnent les fonctions **random** des logiciels de calcul.

Dans le domaine des choix de nombres, c'est-à-dire du hasard *mathématique*, il n'y a pas d'autre forme de hasard que le cryptage du déterminisme. La boutade de Laplace est donc parfaitement juste.

### **I. 5. Hasard et déterminisme dans la Nature.**

Se pourrait-il alors que dans la Nature, au moins, on rencontre du véritable hasard, c'est-à-dire des processus pour lesquels, même si on pouvait les reproduire avec un algorithme déterministe, du moins la Nature n'y aurait pas recouru. Là, le problème est plus délicat, car nous n'avons de la Nature qu'une connaissance approchée, ce qui a pour première conséquence que si deux modèles mathématiques différents donnent des prévisions qui ne diffèrent numériquement que d'une petite quantité, il est impossible de dire que l'un est plus vrai que l'autre, au cas où ces prévisions seraient toutes deux conformes à l'observation. Il est donc absurde de conférer une validité absolue à une description de la Mécanique en termes d'équations différentielles. Par exemple, on peut dire que la bille de la roulette obéit à la Mécanique classique et que son mouvement est décrit par des équations

différentielles, mais cette description n'est qu'un modèle mathématique approché. Le mouvement de la bille est un exemple de chaos déterministe, mais lorsqu'on dit cela il faut avoir en vue que le déterminisme n'existe que dans le modèle mathématique; dans le mouvement réel de la bille, rien ne prouve qu'il ne pourrait être décrit tout aussi bien par un autre modèle mathématique, prédisant le même mouvement pour la bille *dans les limites de précision expérimentales*, mais indéterministe; le déterminisme approximatif du mouvement observé étant alors expliqué comme un effet de la loi des grands nombres et non d'un déterminisme absolu. Or si on parle de modèle indéterministe, on entend par là que le modèle postule un hasard *ontologique*, c'est-à-dire un hasard premier, supposé exister sans explication plus profonde, comme celui qui est invoqué par la Mécanique quantique, par exemple. Mais il suffirait d'ajouter à ce modèle mathématique une annexe "expliquant" le hasard invoqué comme l'application d'un algorithme du type indiqué ci-dessus, pour le transformer en modèle déterministe. De sorte que la question du déterminisme ou de l'indéterminisme *de la Nature* est tout simplement dépourvue de sens logique. Pour que cette question puisse acquérir un sens logique, il faudrait avoir de la Nature une connaissance mathématique rigoureusement exacte et non approchée. L'existence d'innombrables exégèses philosophiques sur le sujet ne suffit pas à transmuter cette question absurde en question sensée (et d'autant moins que la plupart de ces exégèses philosophiques, surtout depuis Emmanuel Kant, s'efforcent précisément de montrer l'absurdité de cette question). On pourra lire à ce sujet Richard P. Feynman *Lectures on Physics*, Tome **3**, **II. 6.**)

Poincaré aurait dit qu'entre les deux modèles il faut choisir le plus simple. Cela ne signifie pas que le plus simple est le plus vrai, mais ne signifie pas non plus que rien n'existe réellement et que tout serait convention. Et cela signifie encore moins que si par hasard le plus simple des deux se trouve être déterministe, on soit fondé à en déduire que la science (en faisant ce choix) aurait prouvé que la Nature est déterministe, ou inversement. Indépendamment de Poincaré, un certain bon sens nous ferait juger très artificiel, ou même farfelu, d'ajouter à un modèle indéterministe une annexe qui ne le rendrait pas plus juste dans ses prédictions, et dont la seule fonction serait de le rendre déterministe. Ce bon sens est simplement le sentiment de la naïveté d'une telle entreprise, et ce sentiment de naïveté provient d'une expérience millénaire, qui nous a toujours montré que lorsqu'on découvre une explication plus profonde à un phénomène, celle-ci est bien plus surprenante et plus retorse que ce qu'on aurait pu imaginer.

Pour comprendre de façon plus précise les conséquences du caractère *approché*, et non mathématique, de la connaissance, nous allons analyser un modèle simplifié de roulette: nous supposons que la bille se déplace

à l'intérieur d'un cercle, mais sur un plan (et non une cuvette); qu'elle rebondit sur le bord du cercle comme une boule de billard; et pour simuler le frottement nous supposons simplement que la vitesse de la bille diminue exponentiellement avec le temps. La description de la vraie roulette serait plus difficile et plus compliquée, mais ne comporterait rien d'essentiellement différent.

Cette roulette simplifiée (ou plutôt, ce *modèle* de roulette simplifiée) est un cas de chaos déterministe intermédiaire entre l'exemple extrêmement simple de la suite  $u_n =$  partie décimale de  $n \cdot e$  et la vraie roulette. En outre, une bille dans un billard circulaire est un exemple qui, contrairement à la suite  $u_n$ , n'est pas uniquement mathématique, car on peut à la fois imaginer une bille matérielle sur un tapis vert, rebondissant sur une bordure circulaire, et le modèle mathématique correspondant. Ce modèle est donc exactement ce qu'il nous faut pour comprendre le hasard dans la Nature. Fondamentalement ce modèle permet non seulement de deviner par analogie ce qui se produit pour la roulette, mais également pour de nombreux phénomènes analogues dans leur principe, tels que par exemple l'agitation thermique, le mouvement brownien, la dissémination de poussières dans l'atmosphère, etc.

La bille est lancée à l'instant 0, avec une vitesse initiale  $v$ . La trajectoire qu'elle va décrire est formée de cordes successives, l'extrémité de chacune étant l'origine de la suivante (voir la figure 4). La réflexion sur le bord obéit bien sûr à la loi de la réflexion, à savoir que la bissectrice de l'angle formé par deux cordes consécutives passe par le centre du cercle. Cet angle reste le même pour toutes les cordes; appelons le  $\theta$ . Si la bordure du billard n'était pas un cercle, cet angle changerait à chaque rebondissement; le problème en serait rendu plus compliqué, mais sans aucun bénéfice pour la discussion.

Si l'angle  $\theta$  est commensurable avec  $\pi$ , c'est-à-dire s'il existe un nombre *rationnel*  $\alpha = p/q$  (quotient de deux entiers  $p$  et  $q$ ) tel que  $\theta = \alpha\pi$ , alors la trajectoire de la bille, formée d'une suite de cordes contiguës, est un polygone régulier: les cordes se superposent périodiquement, et la bille repasse périodiquement sur ses traces. Ce cas particulier est l'analogue de la suite  $v_n =$  partie décimale de  $n \cdot 0.5$ , ou plus généralement  $v_n =$  partie décimale de  $n \cdot \alpha$  (avec  $\alpha$  fractionnaire). Par contre si  $\theta$  est incommensurable avec  $\pi$ , par exemple si  $\pi/\theta = e$ , alors la bille ne repasse jamais deux fois sur la même corde. Si la bille roule très longtemps (si donc le frottement est faible) elle parcourt un très grand nombre de cordes, toutes distinctes, et ces cordes remplissent complètement une couronne (voir figure 4). Remplir complètement signifie que n'importe quelle petite région de la couronne, aussi petite soit-elle, finira tôt ou tard par recevoir la visite de la bille.

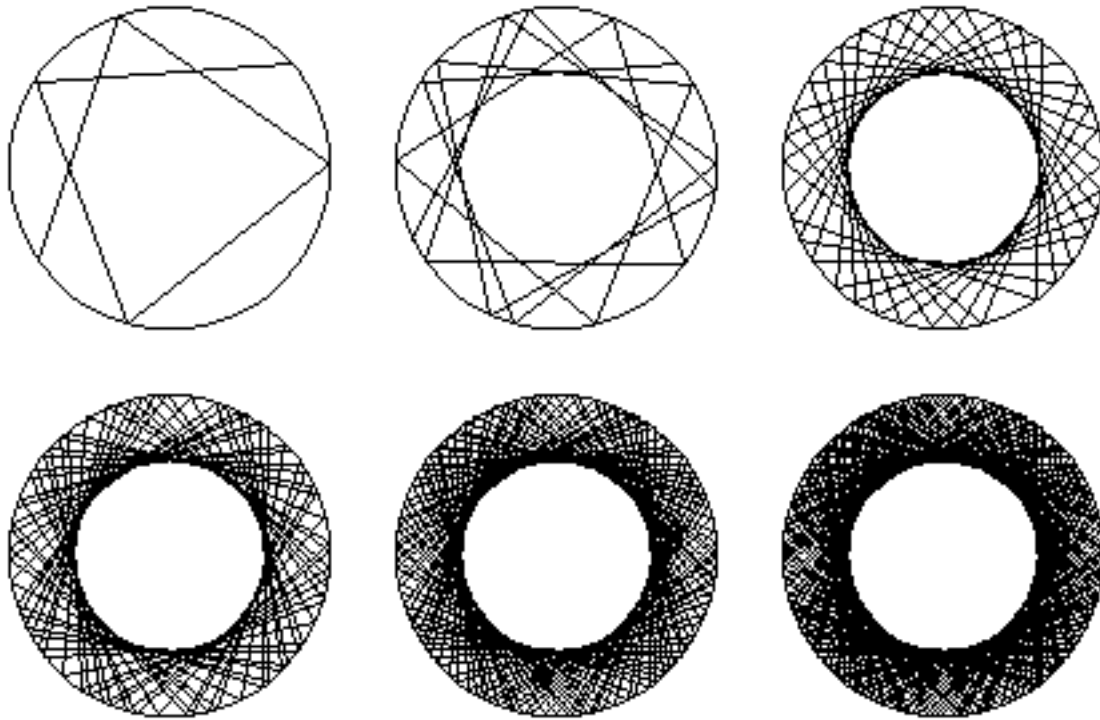


figure 4

Évolution de la trajectoire d'une bille dans un cercle.

Supposons maintenant que le disque soit subdivisé en beaucoup de petites régions. On lance la bille avec une vitesse initiale  $v_0$ , un angle  $\theta = \pi/e$ , et le coefficient de frottement est  $k$ , de sorte que la vitesse de la bille décroît en fonction du temps  $t$  selon la loi  $v(t) = e^{-kt}$ . Les petites régions situées en dehors de la couronne ne seront de toute façon jamais atteintes par la bille. Par contre la bille finira par s'arrêter (plus exactement, elle se rapprochera indéfiniment d'une position limite), forcément dans une région située à une distance du centre supérieure à  $R \sin(\pi/2e)$ . Le modèle est entièrement déterministe, puisque si on connaît  $v_0$ ,  $\theta$ ,  $k$ , et la position initiale  $M_0$  de la bille, alors le point où la bille s'arrêtera est parfaitement déterminé, et la petite région dans laquelle ce point se trouve peut donc être prédite. Voyons comment.

La trajectoire est de toute façon déterminée par la seule donnée de  $M_0$  et  $\theta$ . Le paramètre de frottement  $k$  ne détermine que la distance parcourue sur cette trajectoire. On supposera  $M_0$  situé sur le bord (c'est-à-dire qu'on lance la bille à partir du bord), et en outre l'invariance par rotation permet de ne faire le calcul que pour un point  $M_0$  particulier : si on change  $M_0$  en  $M'_0$  par une rotation, le point d'arrivée sera changé selon la même rotation. La dépendance par rapport à  $M_0$  ne montre donc rien d'intéressant. Mais

supposons qu'on change un tout petit peu  $\theta$  en  $\theta' = \theta + \varepsilon$ . La première corde parcourue par la bille sera légèrement tournée, d'un angle égal à  $\varepsilon$ ; la seconde sera tournée du double, la troisième du triple, etc. Si on appelle  $M_1, M_2, \dots$  les points de la circonférence touchés successivement – après  $M_0$  – par la bille, la distance angulaire de deux consécutifs de ces points est  $\pi - \theta$ , qui se trouve donc modifié en  $\pi - \theta'$ . Étant donné que  $M_0$  est fixé, cela se traduit par un déplacement angulaire de  $-\varepsilon$  pour  $M_1$ , de  $-2\varepsilon$  pour  $M_2$ , de  $-3\varepsilon$  pour  $M_3$ , etc. Au bout d'un nombre  $N$  de réflexions, le point  $M_N$  sera déplacé, par rapport à la trajectoire de référence, de  $-N\varepsilon$ ; si  $N$  est assez grand pour compenser la petitesse de  $\varepsilon$ , la bille pourra se trouver dans une région éloignée de celle qu'elle aurait atteinte après le même nombre de réflexions sur la trajectoire de référence.

Voyons maintenant comment le mouvement dépend de  $k$  (paramètre de frottement). La trajectoire, qu'elle soit de référence ou modifiée, ne change pas lorsqu'on change la valeur de  $k$ ; seule change la *distance parcourue* sur cette trajectoire. Les réflexions successives ne changent rien à la vitesse, seule la direction du mouvement est affectée. On peut donc, bien que la trajectoire soit une ligne brisée, traiter le mouvement *sur cette trajectoire* comme s'il s'agissait d'un mouvement sur un axe. Soit donc  $x(t)$  la distance parcourue *sur cette trajectoire* par la bille à l'instant  $t$ , depuis l'instant 0 où elle était en  $M_0$ . La dérivée de  $x(t)$  est sa vitesse à l'instant  $t$ , dont nous avons supposé qu'elle était égale à  $v_0 e^{-kt}$ . Ainsi :

$$\begin{aligned} x'(t) &= v_0 e^{-kt} \quad ; \\ x(0) &= 0 \quad . \end{aligned}$$

Par un calcul de primitive immédiat, on obtient :

$$x(t) = \frac{v_0}{k}(1 - e^{-kt})$$

On voit que lorsque  $t$  tend vers l'infini,  $x(t)$  tend vers  $v_0/k$ . En pratique  $x(t)$  sera pratiquement égal à  $v_0/k$  déjà pour  $t \sim 10/k$ , et n'en bougera pratiquement plus ensuite, ce qui signifie que la bille s'arrêtera au bout d'un temps de l'ordre de  $10/k$ , après avoir parcouru une distance égale à  $v_0/k$ .

La question est maintenant de savoir en quel lieu du disque la bille aboutira. Pour des valeurs données des paramètres  $k$  et  $\theta$ , la longueur d'une seule corde est  $2R \cos(\theta/2)$  et la distance parcourue  $v_0/k$ , donc le nombre de cordes parcourues (ou, ce qui revient au même, le nombre de rebroussements) est  $N = \text{partie entière de } v_0/2kR \cos(\theta/2)$ , et la position de la bille sur la  $N + 1$ -ième corde sera donnée par la partie décimale de  $v_0/2kR \cos(\theta/2)$  : si cette partie décimale est par exemple 0.75, cela signifiera

que la  $N + 1$  -ième corde aura été parcourue aux trois quarts. Or nous avons vu que si  $\theta$  est modifié d'une très petite quantité  $\varepsilon$ , et que  $N$  est de l'ordre de  $1/\varepsilon$ , la  $N + 1$  -ième corde ne sera pas modifiée un tout petit peu, mais beaucoup. Si  $k$  n'a pas été modifié en même temps que  $\theta$ , la bille, bien qu'ayant parcouru la *même* distance sur la trajectoire, s'arrêtera néanmoins en un point complètement différent du disque. Il suffit pour cela que  $N \sim 1/\varepsilon$ , soit  $v_0/2kR \cos(\theta/2) \sim 1/\varepsilon$  ou, ce qui revient au même :

$$k \sim \frac{v_0 \varepsilon}{2 R \cos(\frac{1}{2}\theta)} \quad (I.3.)$$

D'autre part, si on modifie  $k$  mais pas  $\theta$ , mettons qu'on augmente ou diminue  $k$  d'une proportion  $\eta$ , qu'on le remplace donc par  $k' = k(1 + \eta)$ , alors la trajectoire reste inchangée, mais la distance parcourue sur cette trajectoire passera de  $v_0/k$  à  $v_0/k'$ , c'est-à-dire qu'elle sera divisée par  $1 + \eta$ , ou multipliée par  $1 - \eta$  puisque pour  $\eta$  petit  $1/(1 + \eta) \simeq 1 - \eta$ . La différence entre les deux est donc à peu près égale à  $\eta v_0/k$ , ce qui n'est pas petit si  $k$  est du même ordre de grandeur que  $\eta v_0$ . Cela peut même dépasser la longueur d'une corde si

$$k < \frac{v_0 \eta}{2 R \cos(\frac{1}{2}\theta)} \quad (I.4.)$$

On voit alors ce qui se passe: tant que le problème reste purement mathématique, la position finale est toujours exactement prédictible à partir de la donnée des paramètres  $\theta$  et  $k$ . Mais d'après (I.3) et (I.4.) l'incertitude sur la position finale est beaucoup plus grande que l'incertitude sur les valeurs de ces paramètres, et cela d'autant plus que  $k$  est plus petit. Autrement dit : plus le frottement est faible, plus cette disproportion entre les incertitudes est grande. Pour en donner une idée numérique : supposons que le temps soit mesuré en secondes et les distances en décimètres, que le rayon du disque soit  $R = 1$  décimètre, et  $v_0 = 1$  décimètre par seconde. Prenons  $k = 10^{-3}$ ,  $\theta$  tel que  $\cos(\theta/2) = 0.75$ , ce qui fait  $\theta \simeq 1.445468496$  et  $\pi/\theta \simeq 2.173407904$ . La longueur d'une corde sera alors 1.5, la distance parcourue par la bille  $10^3$ , c'est-à-dire 100 mètres, le nombre de cordes parcourues sera  $N = 666$ , et la bille s'arrêtera aux deux tiers de la 667<sup>e</sup> corde. On peut se donner un repère orthonormé pour les coordonnées des points, tel que  $M_0$  ait pour coordonnées (1, 0). Puisqu'on passe de chaque corde à la suivante par une rotation d'angle  $\pi - \theta$ , on aura la 667<sup>e</sup> corde par une rotation d'angle  $666 \cdot (\pi - \theta) \simeq 1129.618689203$ , ce qui est égal à  $4.928519218$  modulo  $2\pi$ . Le point situé aux deux tiers de cette corde est alors le point de coordonnées  $x = +0.699662113$  et  $y = -0.102337320$ . Si on prend  $\varepsilon = 10^{-3}$ , et  $\eta = 0$ , le nouvel angle  $\theta'/2$  a pour cosinus non plus 0.75, mais 0.749669187, la longueur des cordes n'est plus 1.5, mais 1.499338375, le

nombre de cordes parcourues est toujours 666, mais la proportion parcourue de la 667<sup>e</sup> est 0.960 852 116. Si on refait les mêmes calculs de rotations, on trouve que cette fois le point atteint par la bille est le point de coordonnées  $x = +0.893\ 362\ 414$  et  $y = -0.342\ 554\ 750$ . Les deux points sont à l'intérieur du disque, mais leur distance mutuelle est 0.308 584 218, à peu près le tiers du rayon. Le calcul confirme donc ce qui était qualitativement prévu, à savoir qu'un changement de un millième de l'angle  $\theta$  entraîne un changement macroscopique du point atteint par la bille. On observerait la même chose si au lieu de faire varier  $\theta$ , on avait fait varier  $k$ , en prenant  $\varepsilon = 0$  et  $\eta = 10^{-3}$ , ou en faisant varier les deux à la fois. Cela veut dire que si on avait voulu prédire le point où la bille viendra s'immobiliser au millième près, il aurait fallu connaître  $\theta$  et  $k$  avec une précision non du millième, mais du millionième. C'est ce que montrent les relations (I.3.) et (I.4.). Le point où la bille s'arrêtera peut donc être prédit, mais la précision de cette prédiction exige une précision encore bien plus grande sur les valeurs des paramètres initiaux  $\theta$  et  $k$ , d'autant plus grande d'ailleurs que  $k$  est plus petit, conformément à (I.3.). Il s'agit là d'une propriété caractéristique du chaos déterministe, qui est l'*amplification* des perturbations, ou la *très haute sensibilité aux conditions initiales*.

D'autre part, il se produit – mais en deux dimensions cette fois – le même phénomène que pour la suite  $u_n$ , à savoir que si le frottement était absolument nul, la bille ne s'arrêterait jamais, mais parcourrait la couronne éternellement, et si on délimitait de petites régions dans la couronne, alors

a) chacune de ces petites régions serait traversée épisodiquement par la bille (d'autant plus souvent que la durée de l'observation serait longue);

b) le temps total (formé donc de nombreuses traversées brèves) que la bille passerait à l'intérieur de chacune de ces petites régions, dépendrait de l'aire de cette petite région.

Pour un être microscopique vivant dans l'une de ces petites régions, et voyant de temps en temps passer la bille, les passages de cette bille sembleraient survenir à des moments aléatoires (cela provient de ce que  $\pi/\theta$  est irrationnel). L'être microscopique, ne pouvant pas imaginer la cause déterministe des passages de la bille, penserait avoir affaire à un hasard ontologique. En cas de frottement, il s'apercevrait cependant que la vitesse de la bille est plus lente à chaque passage, et il pourrait même avoir le privilège fantastique de voir un jour la bille venir s'arrêter dans son jardin.

À une toute autre échelle, et avec une perspective complètement différente, le physicien humain qui lance cette bille serait un dieu pour l'être microscopique, mais, sa connaissance des choses étant approchée et non mathématiquement exacte, il ne pourrait lui-même prévoir le point où

la bille s'arrêtera, faute de pouvoir mesurer avec une précision suffisante les conditions initiales du mouvement. Ainsi, le postulat selon lequel le mouvement est déterministe a un sens dans le modèle mathématique, mais n'en a absolument aucun dans la réalité.

Une question intéressante est maintenant de savoir comment sont répartis ces temps de séjour dans chaque petite région. Avec la suite  $u_n$ , nous avons dit que (pour  $n$  grand) le nombre de termes qui "tombent" dans un intervalle  $[a, b]$  de  $[0, 1]$  est proportionnel à la longueur  $b - a$ . Pour la bille, nous allons voir que la durée totale de séjour dans chacune des petites régions n'est pas proportionnelle à son aire : les régions situées près du bord intérieur de la couronne seraient – à aire égale – fortement favorisées par rapport aux régions situées près du bord extérieur. On peut comprendre cela facilement : ce qui est réparti uniformément n'est pas la densité des passages de la bille, mais les angles d'inclinaison des cordes : en effet, ces angles sont, comme nous l'avons vu, en progression arithmétique, et la suite des angles, comptés modulo  $2\pi$ , est donc du même type que la suite  $u_n$ . Si on se fixe deux cordes proches l'une de l'autre, leurs angles d'inclinaison sont proches également, ils diffèrent, disons, d'un petit angle  $\delta$  (figure 5). Cela veut dire que ces deux cordes ont entre elles un angle  $\delta$  ; elles se coupent tout près du bord intérieur de la couronne, mais s'écartent l'une de l'autre en s'éloignant de l'intersection. Parmi toutes les cordes de la trajectoire, certaines se situeront entre les deux cordes que nous nous sommes fixées ; leur nombre sera proportionnel à l'angle  $\delta$ , tout comme le nombre d'éléments de la suite  $u_n$  situés à l'intérieur d'un petit intervalle de longueur  $\delta$  est proportionnel à  $\delta$ . Le nombre de ces cordes sera à peu près égal à  $\delta/2\pi$  fois le nombre de toutes les cordes parcourues. Or les cordes sont plus denses, ou plus serrées, près du bord intérieur de la couronne : leur distance est nulle tout près du bord intérieur et croît ensuite linéairement, ce qui entraîne que leur densité est très grande près du bord intérieur (à la limite infinie), mais diminue en se rapprochant du bord extérieur. Cela se voit immédiatement rien qu'en regardant la figure 4.

Pour calculer, introduisons la coordonnée radiale  $t = \sqrt{x^2 + y^2}/R$  (la distance au centre du disque rapportée au rayon). La couronne correspond aux valeurs de  $t$  comprises entre  $r = \sin(\frac{1}{2}\theta)$  et 1 ( $rR$  est ainsi le rayon intérieur de la couronne).

Les aires infinitésimales  $\sigma$  délimitées par deux couples de cordes très proches, formant entre elles de petits angles  $\delta$  (voir figure 5) sont égales à

$$\sigma = \frac{1}{2} t^2 \delta^2 \sqrt{\frac{t^2}{r^2} - 1}$$

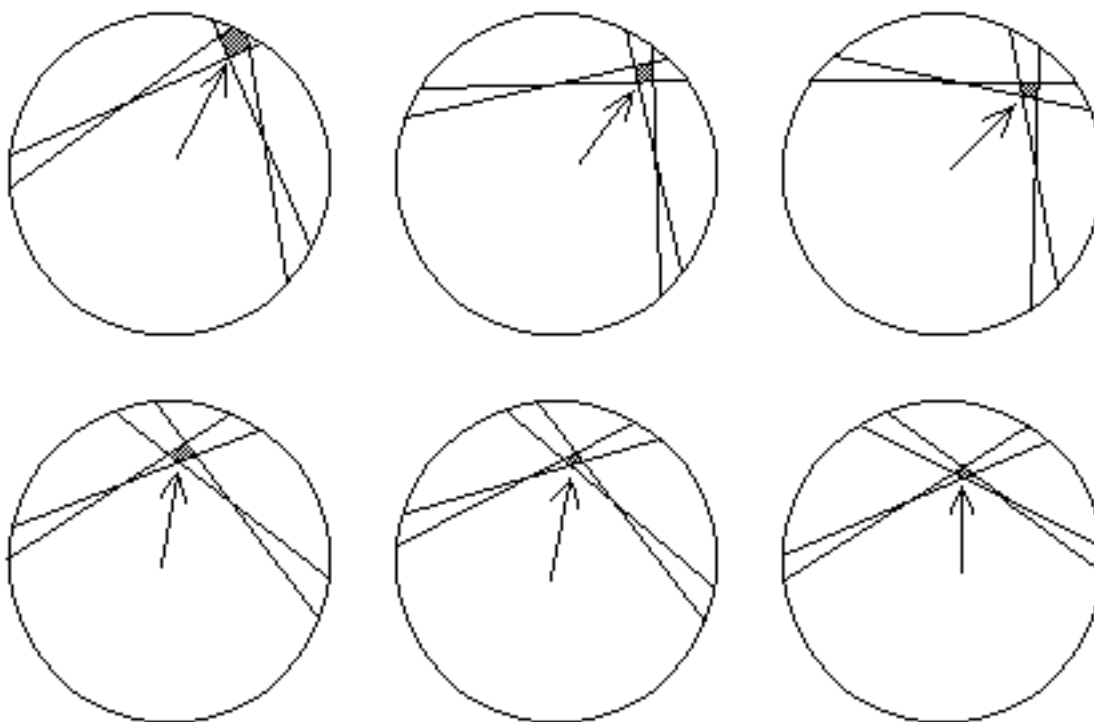


figure 5

Petits losanges délimités par des faisceaux de cordes.

Mais puisque, entre deux cordes fixées, ayant entre elles un angle  $\delta$ , le nombre de passages de la bille est proportionnel à  $\delta$  (plus précisément égal à  $N\delta/2\pi$ ,  $N$  étant le nombre total de cordes parcourues), et le petit intervalle que la bille parcourt à chaque passage à l'intérieur du losange étant de longueur  $t^2\delta/2r$ , on en déduit que la distance totale (somme de tous ces petits intervalles) parcourue par la bille dans le petit losange d'aire  $\sigma$  est

$$\tau = \frac{N}{2\pi r} t^2 \delta^2$$

Si on exprime  $\tau$  en fonction de  $\sigma$  on trouve

$$\tau = \frac{N}{\pi \sqrt{t^2 - r^2}} \sigma$$

Conformément à ce que nous avons pressenti auparavant par des estimations qualitatives,  $\tau$  n'est pas simplement proportionnel à  $\sigma$ ; on voit apparaître la densité  $1/\sqrt{t^2 - r^2}$ , qui exprime quantitativement le favoritisme dont bénéficient, à aire égale, les petits losanges situés près du bord intérieur de la couronne. Cette loi valable pour les petits losanges s'étend ensuite à des petites régions de forme quelconque, puisqu'on peut toujours quadriller

une région selon de tels petits losanges, comme on quadrille en petits carrés dans la théorie usuelle de l'intégration.

Si au lieu de compter le *chemin* parcouru à l'intérieur de chaque petit losange on avait voulu compter simplement le *nombre de passages*, sans considération du chemin parcouru, on aurait aussi obtenu une loi statistique, dont on aurait déduit une probabilité empirique; dans ce cas, ce seraient les dimensions linéaires, et non plus l'aire, qui interviendraient: par exemple, si la petite région est un petit disque de diamètre  $\rho$ , le nombre de passages de la bille à travers ce petit disque serait à peu près égal à

$$\nu = \frac{N}{\pi \sqrt{t^2 - r^2}} \rho$$

Bien entendu la densité  $1/\sqrt{t^2 - r^2}$  est toujours là.

Ainsi les êtres microscopiques qui habitent dans la couronne observeraient des lois statistiques, vérifiées par les passages de la bille: pour toutes les régions situées à *une distance donnée* du centre du disque, le total du temps passé par la boule à l'intérieur d'une région du disque est – sur un grand nombre de passages – proportionnel à l'aire de cette région; ou encore: le nombre total de passages à travers une petite région est proportionnel à son diamètre. Mais lorsqu'on considère des régions situées à des distances différentes du centre, le coefficient de proportionnalité varie selon  $1/\sqrt{t^2 - r^2}$ .

En observant vraiment bien, les êtres microscopiques apercevraient des régularités dans les passages, car si la suite des passages semble aléatoire à première vue, elle a les mêmes défauts que la suite  $u_n$ . Ceci provient de ce que le modèle est trop simple, on peut tout calculer explicitement à l'aide de formules ou d'algorithmes simples, que des inspecteurs peuvent décrypter. Mais on peut rendre le modèle plus compliqué: par exemple, au lieu d'une bordure circulaire, on pourrait prendre une bordure non circulaire, pas même elliptique (car l'ellipse possède encore trop de régularités). Point n'est besoin de modifier beaucoup la bordure: celle-ci peut ne jamais s'écarter du cercle de plus de un centième de millimètre: nous avons vu qu'en ne changeant l'angle  $\theta$  que d'un millième, on perturbait macroscopiquement la forme prise par la trajectoire après plusieurs centaines de rebroussements; il en irait de même en modifiant la bordure de quelques millièmes; en effet, modifier cette bordure revient à modifier *tous* les angles de réflexion, chacun d'une quantité petite, mais différente, et non pas seulement le premier. En agissant ainsi, on produirait le même type de perturbation amplifiée: le début de la trajectoire serait peu modifié, mais *après un certain temps*, inversement proportionnel à l'amplitude des modifications,

elle serait complètement différente. Le simple fait d'avoir une bordure qui, bien que presque circulaire, n'est plus rigoureusement un cercle, rend déjà le calcul exact de la trajectoire si compliqué que les inspecteurs n'y verraient que du feu. Pour déjouer encore davantage leur vigilance, on pourrait donner au fond du disque (le tapis vert) une forme courbe de cuvette au lieu d'un plan. Le mouvement serait toujours rigoureusement déterministe. Les propriétés de chaos que nous avons mises en évidence dans le modèle simple subsisteraient (à savoir l'*amplification*, le fait qu'une très faible modification des conditions initiales entraînerait une modification macroscopique du mouvement après un nombre de tours inversement proportionnel à cette modification). Mais l'algorithme pour les calculer serait alors tellement complexe que ce mouvement *ne pourrait pas être calculé*. Les êtres microscopiques ne pourraient plus observer des régularités dans les passages individuels de la bille : si leurs observations étaient exactes (c'est-à-dire sans erreur de mesure), ce serait la complexité du processus qui les empêcherait de voir les régularités. Et pour être bien certain que les habitants de la couronne, qu'il ne faudrait tout de même pas sous-estimer, ne pourront jamais percer le mystère des passages de la bille, nous supposons que la zone située près du bord leur est inaccessible (que les lois physiques de leur monde sont telles qu'il leur faudrait une énergie infinie pour atteindre ce bord) et qu'ils ne peuvent donc pas connaître exactement sa forme. Si en outre leurs observations sont approchées et entachées d'erreurs de mesure, le déterminisme du mouvement leur sera encore plus caché. Les seules propriétés suffisamment simples qu'ils pourront observer seront les deux propriétés suivantes :

a) la durée totale de séjour dans une petite région d'aire  $\sigma$  située à une distance  $t$  du centre du disque est proportionnelle à  $\sigma$  et inversement proportionnelle à  $\sqrt{t^2 - r^2}$ .

b) la variation en  $1/\sqrt{t^2 - r^2}$  s'explique *simplement* par une propriété géométrique. C'est-à-dire que les êtres microscopiques peuvent aisément comprendre que les trajectoires sont des segments de droites qui s'écartent linéairement et sont donc moins denses vers le bord.

Ainsi le phénomène se décompose en deux parties : 1. une distribution mystérieuse des segments de trajectoire, et 2. un effet dû uniquement aux propriétés géométriques communes à tous les segments (théorème de Thalès, etc.). La partie 2 du phénomène est compréhensible et "explique" la densité  $1/\sqrt{t^2 - r^2}$ . Par contre, la partie du phénomène que les êtres microscopiques *ne peuvent pas comprendre* est la distribution des segments. Il leur est aisé de voir que les segments se coupent près du bord intérieur (comme on voit sur la figure 5) et utiliser le théorème de Thalès pour retrouver par le calcul la densité  $1/\sqrt{t^2 - r^2}$ , mais ils ne peuvent pas prévoir la position des segments successifs. Ils peuvent alors découvrir après quelques

tâtonnements mathématiques que, si au lieu de mesurer le temps de séjour dans chaque petite région, on avait mesuré par exemple l'orientation des segments (l'angle d'inclinaison par rapport à la direction radiale), alors il n'y aurait plus de densité variable : les différentes orientations se distribueraient de façon uniforme. Étant alors donné que la distribution des segments est incompréhensible, il n'est plus possible de progresser au-delà de cette loi uniforme : de même que dans un calcul algébrique (par exemple factoriser un polynôme ou réduire une fraction du type  $(\sqrt{2} + \sqrt{3})/(\sqrt{5} - \sqrt{3})$  à sa forme la plus simple), il arrive forcément un moment où "on ne peut plus réduire davantage", de même lorsqu'un phénomène est décomposé en une partie incompréhensible et une partie compréhensible, il arrive forcément un moment où "on ne peut plus réduire davantage" la partie incompréhensible. On a alors atteint un stade de *compréhension maximum* du phénomène. À un tel stade, la partie incompréhensible est réduite à une distribution aléatoire uniforme, car tant qu'il subsiste une non-uniformité, celle-ci peut faire l'objet de recherches pour en comprendre l'origine, comme ce fut le cas pour la densité  $1/\sqrt{t^2 - r^2}$ . En revanche, une fois la partie incompréhensible réduite "à sa plus simple expression", le seul progrès encore possible consisterait à lever l'obstacle du brouillage par le chaos (par exemple disposer d'ordinateurs tellement puissants et d'instruments de mesure tellement précis que le chaos déterministe puisse être lui aussi maîtrisé).

Une fois qu'on a regroupé tout ce qui est compréhensible et réduit l'incompréhensible à sa plus simple expression, uniforme par nature, ce dernier reçoit alors un nom : le hasard.