Marches aléatoires

Bruno Aebischer* d'après Emmanuel Cepa

Vendredi 17 novembre 2006

Table des matières

* bruno.aebischer@univ-fcomte.fr

I. Présentation du problème

I.1. Origine de cet exposé

Cet exposé fait suite à une rencontre avec Emmanuel Cépa (de l'Université d'Orléans, et nouveau directeur de l'IREM de cette Université) dans le cadre de la CIIU (Commission Inter IREM Universitaire). Lors de cette réunion de la CIIU, Emmanuel nous a présenté une conférence qu'il avait faite à des lycéens pour leur présenter une démarche de recherche mathématique. Il avait donc essayé de choisir un domaine (proche de ses compétences : c'est un probabiliste) pour lequel très peu de prérequis étaient nécessaires, et auquel on pouvait trouver beaucoup de débouchés. C'est ainsi qu'il avait choisi de parler de dénombrement de chemins, qui débouche naturellement sur un problème de marches aléatoires mais peut s'appliquer aussi dans d'autres situations diverses : observation du dépouillement d'un scrutin électoral, choix d'un fond de caisse pour un organisateur de spectacle...
De plus, nous avions écoutés récemment Sébastien Darses nous parler de mouvements Browniens, et les problèmes de marches aléatoires sont étroitement liés aux mouvements Browniens ! C'est donc une raison de plus qui m'a poussé à proposer ce petit exposé.

I.2. Prérequis

Pour comprendre cet exposé, un lycéen a besoin d'un tout petit bagage ensembliste et probabiliste intuitif :
Admettons qu'événements et ensembles peuvent se comprendre sans les définir.

La probabilité d'un événement se définit et se comprend ainsi :

$$\mathrm{P}(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

On utilisera les propriétés suivantes :

- Un événement certain admet 1 comme probabilité ;
- Si A et B sont deux événements incompatibles, $\mathrm{P}(A \sqcup B) = \mathrm{P}(A) + \mathrm{P}(B)$ [1] ;
- $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$

Ces deux dernières propriétés étant en fait une conséquence immédiate d'un principe de dénombrement élémentaire et intuitivement évident : si A et B n'ont pas d'élément en commun, si C désigne l'ensemble des éléments qui sont dans A ou dans B, (on note $C = A \cup B$ ou plus précisément, justement dans ce cas $C = A \sqcup B$, car ici le « ou » est exclusif), alors le nombre d'éléments de C est bien sûr la somme du nombre d'éléments de A et du nombre d'éléments de B. En notant $\mathrm{card}(A)$ le nombre d'éléments de l'ensemble A, on a donc $\mathrm{P}(A \sqcup B) = \mathrm{P}(A) + \mathrm{P}(B)$ lorsque A et B n'ont pas d'élément en commun.

II. Exposés des problèmes abordés

Bien sûr ces problèmes « pseudo-concrets » sont simplifiés à l'extrême, et les situations concrètes risquent d'être beaucoup plus compliqués !

[1] en notant $A \sqcup B$ la réunion de A et B dans le cas où ces événements sont incompatibles, c'est-à-dire, d'un point de vue ensembliste, lorsque leur intersection est vide

II.1. Dépouillement d'une élection

Au cours d'un scrutin opposant deux candidats, A obtient 600 voix tandis que 400 voix se sont portées sur B (ne tenons pas compte des bulletins nuls)...
En fait, le dépouillement n'a pas encore eu lieu, et personne ne connaît, pour l'instant le résultat. Les bulletins sont dépouillés un par un, et les décomptes faits au fur et à mesure. Quelle est la probabilité pour qu'au cours du dépouillement, A ait été tout le temps majoritaire (au sens large) ? (On pourrait ensuite essayer d'inverser le problème de probabilités et essayer d'estimer le résultat final après un dépouillement partiel, mais ce serait toute une autre histoire de statistiques inférentielles).

II.2. Fond de caisse

Un organisateur de spectacle attend 200 spectateurs. Le prix du spectacle est 5€ par personne. On admet que parmi les 200 spectateurs, 100 ont prévu un billet de 5€ pour payer, alors que les 100 autres ont un billet de 10€ et il faudra leur rendre la monnaie. Les spectateurs avec ou sans l'appoint se présentent dans un ordre aléatoire. Quel fond de caisse l'organisateur doit-il prévoir pour avoir moins de 5% de chances de tomber à cours de monnaie ?

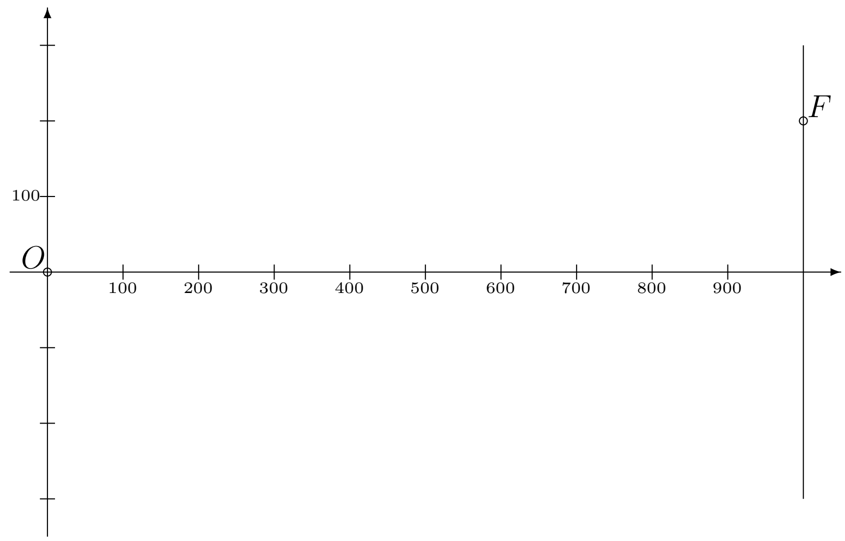
II.3. Marche aléatoire

Un marcheur (sans doute ivre !) part du point d'abscisse 0 (sa maison), et se promène au hasard dans sa rue (infinie !). Après chaque pas (l'unité est le pas !) il lance une pièce pour savoir s'il fait le pas suivant en avant (si c'est pile) ou en arrière (face). Quelle est la probabilité que le marcheur revienne un jour devant sa maison ?
Et que se passe-t-il dans le cas d'une marche aléatoire dans un quadrillage avec un dé à quatre faces, et à chaque nœud, un arrêt et un nouveau choix de la direction de continuation vers le Nord, l'Ouest, le Sud ou l'Est, décidé par le lancement du dé ?
Et pour un oiseau ivre qui peut en plus monter ou descendre (et qui utilise donc un dé à 6 faces !)

Ce dernier problème (quelle qu'en soit sa dimension) est une approche du mouvement Brownien pas si farfelue que ça : on peut montrer que c'en est une approximation réaliste.

III. Modélisation commune

III.1. Dépouillement



Considérons dans un plan muni d'un repère les points de coordonnées entières. On placera un point par bulletin dépouillé de la manière suivante : son abscisse représente le numéro d'ordre du bulletin dépouillé, et en ordonnée la différence entre le total partiel des bulletins A et le total partiel des bulletins B. Le premier point est bien sûr $O(0, 0)$, et le dernier est $F(1000, 200)$. On joint ces 1001 points par des segments de droites, et on obtient ainsi un chemin menant de A à B. Deux points d'abscisse consécutive ont une ordonnée qui diffère d'une unité, le chemin monte ou descend selon que le dernier bulletin dépouillé est pour A ou pour B.
A aura toujours été en tête lorsque le chemin qu'on représente est toujours au-dessus (au sens large) de l'axe des abscisses, et la probabilité que A soit toujours en tête du dépouillement est égale au quotient du nombre de chemins ne passant pas en dessous de l'axe des abscisses par le nombre total de chemins joignant O à F.

III.2. Fond de caisse

Dans le même quadrillage que précédemment, plaçons en abscisse le numéro de passage à la caisse du spectateur et en ordonnée le nombre de billets de 5€ qui restent en caisse. Ici encore, les ordonnées de deux points d'abscisse consécutive diffèrent de 1, et un passage possible des spectateurs est modélisé par un chemin qui part de $A(0, n)$ et qui arrive à $B(400, n)$, n étant le nombre de billets dans le fond de caisse.
De la même manière, l'organisateur tombera à cours de monnaie si le chemin passe en dessous de l'axe des abscisses.

III.3. Marche aléatoire

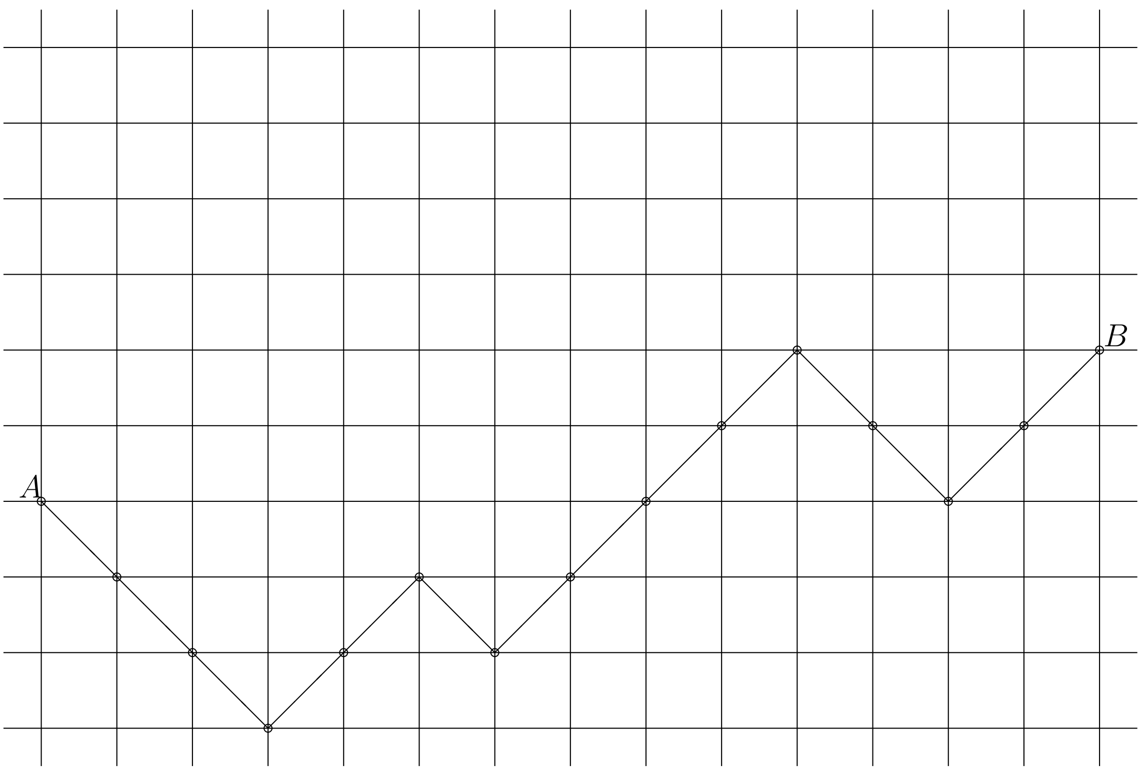
Dans la même grille on placera en abscisse le nombre de pas que le marcheur aléatoire a fait, et en ordonnée la distance algébrique qu'il a parcourue depuis sa maison. Ici aussi, un résultat de la marche aléatoire peut donc être modélisée exactement comme un chemin qui part de l'origine, avec comme contrainte d'aller toujours vers la droite, et de monter ou de descendre d'une unité à chaque pas.
Le problème posé est un peu plus compliqué, car on ne peut pas compter tous les chemins possibles (la longueur n'étant pas fixée, il y en a une infinité). On calculera néanmoins la probabilité pour

qu'au bout de $2n$ pas le marcheur soit revenu à son point de départ, et on fera un passage à la limite.

IV. Dénombrement de chemins

IV.1. Définition des chemins

On considère un quadrillage régulier du plan. Les chemins que l'on va considérer doivent joindre des points du quadrillage en allant de gauche à droite uniquement, et en suivant toujours des diagonales du quadrillage :



Voici un exemple de chemin joignant A et B. Nous allons essayer de dénombrer tous les chemins de ce type joignant deux points du plan.

Il est clair qu'un chemin est caractérisé par son point de départ (ici A), sa longueur ℓ (ici ce chemin a pour longueur 14) et par une succession de ℓ montées (m) ou descentes (d). Le point de départ étant fixé (tous les points du plan quadrillé supposé infini sont bien sûr équivalents), un chemin peut donc être aussi modélisé à son tour par un mot formé des deux lettres m et d. Par exemple le chemin représenté ci-dessus est modélisé par $dddmmdmmmmddmm$.

Si on sait que le chemin part de A, il arrivera à B si sa longueur est 14 et s'il comporte deux montées de plus que de descentes, donc s'il contient 8 fois la lettre m et 6 fois la lettre d. Il y a donc autant de chemins joignant A et B que de mots de 14 lettres formés de 8 m et de 6 d.

Pour définir un tel mot, il suffit de définir l'emplacement des 8 m, il suffit donc de choisir un sous-ensemble de 8 éléments parmi les 14 premiers nombres. Par exemple, le chemin précédent correspond au sous-ensemble $\{4, 5, 7, 8, 9, 10, 13, 14\}$. Il y a donc autant de chemins menant de A à B que de parties de 8 éléments dans l'ensemble des 14 premiers nombres. Ce nombre se note $\binom{14}{8}$ (on lit 8 parmi 14). [2]

[2] On pourrait aussi bien raisonner en commençant par définir l'emplacement des six descentes, et on trouverait qu'il y a donc $\binom{14}{6}$ chemins de A vers B, donc on prouve ainsi que $\binom{14}{6} = \binom{14}{8}$

IV.2. Généralisation

Soient A et B deux points quelconques du plan. On suppose juste que B est à droite de A, c'est-à-dire que son abscisse est plus grande que celle de A. Soit Δx la différence d'abscisse entre A et B, et soit Δy leur différence d'ordonnées. Dans l'exemple précédent, $\Delta x = 14$ et $\Delta y = 2$. D'après les hypothèses du problème, Δx et Δy sont deux entiers, et on est sûr que $\Delta x > 0$, alors que Δy peut être positif, nul ou négatif.
Lorsqu'on fait p montées et q descentes en partant de A, on n'atteindra B que si on a $p+q = \Delta x$ et $p-q = \Delta y$. On peut alors exprimer p et q en fonction de Δx et Δy : on a bien sûr $p = \frac{1}{2}(\Delta x+\Delta y)$ (et $q = \frac{1}{2}(\Delta x - \Delta y)$). En reprenant le raisonnement qu'on a fait sur l'exemple, un chemin de longueur Δx mène bien de A à B s'il est formé de $p = \frac{1}{2}(\Delta x + \Delta y)$ montées (et donc de $q = \frac{1}{2}(\Delta x - \Delta y)$) descentes, donc s'il est modélisé par un mot de Δx lettres m ou d, comprenant p fois la lettre m (et donc q fois la lettre d), et il y a donc autant de chemins de A vers B que de sous-ensembles de p éléments parmi les Δx premiers nombres.
Il y a donc dans ce cas $\binom{\Delta x}{\frac{\Delta x+\Delta y}{2}}$ chemins menant de A vers B [3].
Il est évidents que p doit être un entier, donc que Δx et Δy doit être de même parité, et aussi qu'on doit avoir $0 \leqslant p \leqslant \Delta x$. Nous énonçons ce premier résultat.

IV.3. Premiers résultats

Définition 1 *Deux points du quadrillage sont dits* joignables *lorsqu'il existe au moins un chemin qui les relie.*
Un point est accessible (sous-entendu : accessible depuis l'origine) lorsqu'il existe un chemin entre le point $O(0,0)$ et lui.

Bien entendu, il est tout à fait équivalent de partir de O ou d'un autre point A, à condition d'appliquer à l'extrémité la même translation que celle qui ramène en O le point A.

Théorème 1 *Deux points de coordonnées entières A et B sont joignables si et seulement si la différence de leurs abscisse Δx et la différence de leurs ordonnées Δy vérifient :*
- $\Delta x \in \mathbb{N}$ *;*
- Δy *a la même parité que* Δx *;*
- $-\Delta x \leqslant \Delta y \leqslant \Delta x$

Un point $A(a,b)$ est accessible si et seulement si ses coordonnées vérifient les trois conditions suivantes :
- $(a,b) \in \mathbb{N} \times \mathbb{Z}$ *(a et b sont entiers, a est positif ou nul).*
- $a+b$ *est pair. (a et b ont la même parité).*
- $|b| \leqslant a$

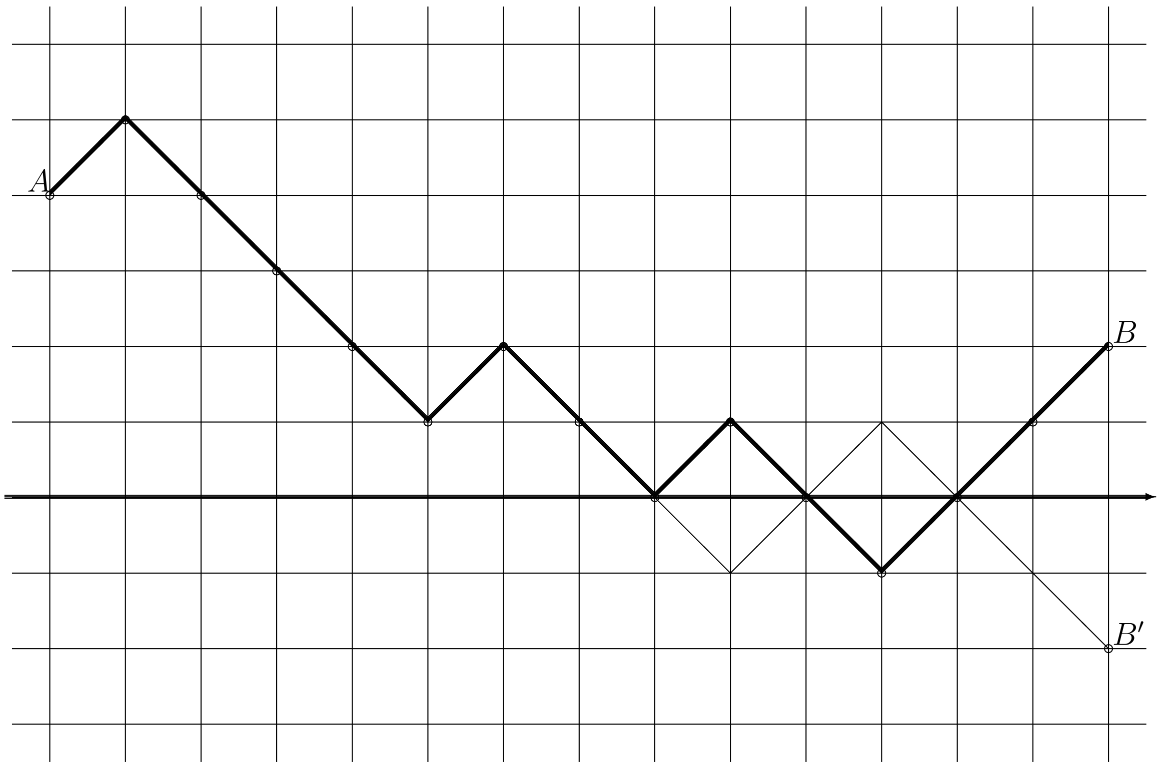
IV.4. Principe de réflexion

Voici maintenant un résultat très important pour résoudre le problème du marcheur.

Théorème 2
Soit A et B deux points joignables situés du même côté de l'axe des abscisses.
Soit B' le point symétrique de B par rapport à l'axe des abscisses.
Alors le nombre de chemins qui vont de A vers B et qui touchent l'axe des abscisses est égal au nombre de chemins allant de A vers B'.

[3] En raisonnant sur les descentes, on montre que ce nombre est aussi égal à $\binom{\Delta x}{\frac{\Delta x-\Delta y}{2}}$

Démonstration :
Soit un chemin menant de A à B et qui touche l'axe des abscisses. On considère le chemin obtenu en conservant la partie situées *avant le premier point de contact*, et en symétrisant par rapport à l'axe des abscisses sa partie située *après le premier point de contact.*
On obtient un chemin menant de A à B'.
Réciproquement, un point menant de A à B' coupe forcément au moins une fois l'axe des abscisses. En symétrisant la partie de ce chemin située à droite du premier point de contact, on rejoint le point B. On définit une correspondance *biunivoque* (les anglophones diraient "one to one") entre les chemins allant de A à B touchant l'axe des abscisses et tous ceux qui vont de A vers B', qui prouve qu'il y a autant de chemins dans les deux catégories.



IV.5. Notations

Avant d'attaquer le cœur du problème, nous allons définir quelques notations utiles.
Soient A et B deux points joignables. On notera $\mathrm{nch}(AB)$ le nombre de chemins joignant A et B.
On notera $[AB]$ l'ensemble de tous les chemins joignant A et B.
On a donc $\mathrm{nch}(AB) = \mathrm{card}([AB])$.
Si C est un autre point, $[ACB]$ désignera l'ensemble des chemins joignant A à B en passant par C, on pourra ainsi ajouter d'autres points fixes pour décrire un ensemble de chemins.
Si A et B sont situés du même côté de l'axe des abscisses, on notera $[\widehat{AB}]$ l'ensemble des chemins joignant A et B sans rencontrer l'axe des abscisses.
$\mathrm{nch}(\widehat{AB})$ est le nombre de chemins joignant A et B sans rencontrer l'axe des abscisses :

$$\mathrm{nch}(\widehat{AB}) = \mathrm{card}([\widehat{AB}])$$

Avec ces notation, le résultat du principe de réflexion peut se traduire par

$$\mathrm{nch}(\widehat{AB}) = \mathrm{nch}(AB) - \mathrm{nch}(AB').$$

Enfin, si M et N sont deux points de l'axe des abscisses, $[\widetilde{MN}]$ désignera l'ensemble des chemins joignant M et N, qui ne rencontrent pas l'axe des abscisses à part en M et en N, et le nombre de tels chemins sera noté $\mathrm{nch}(\widetilde{MN}) = \mathrm{card}[\widetilde{MN}]$.
Enfin, $\sqcup$ est le symbole que nous utiliserons pour désigner la réunion de deux ensembles disjoints. On a bien sûr $\mathrm{card}(A \sqcup B) = \mathrm{card}(A) + \mathrm{card}(B)$.

IV.6. Théorème d'itération

Nous sommes maintenant armés pour énoncer le résultat essentiel qui nous permettra de résoudre le problème des marches aléatoires.

Théorème 3 *Soit $O(0,0)$. Soient N et N' deux point accessibles consécutifs de l'axe des abscisses : $N(2n-2,0)$ et $N'(2n,0)$.*
Alors $\mathrm{nch}(\widetilde{ON'}) = 4\,\mathrm{nch}(ON) - \mathrm{nch}(ON')$.

Ce théorème permet donc d'exprimer le nombre de chemins joignant un point accessible de l'axe des abscisses sans toucher cet axe *en dehors de ses extrémités en fonction tout simplement de deux nombres de chemins.*
Démonstration :
Soit $A(1,1)$ et $A'(1,-1)$; Soit $B(2n-1,1)$ et $B' = (2n-1,-1)$.

Un chemin joignant O à N' doit nécessairement passer par A ou par A', et aussi par B ou par B'.
D'autre part, si un chemin joignant O à N' ne rencontre pas l'axe des abscisses ailleurs qu'en O et N', c'est que entre A et B (ou entre A' et B') il ne rencontre pas l'axe des abscisses. On peut noter cette remarque ainsi :
$[\widetilde{ON'}] = [O\widehat{AB}N'] \sqcup [O\widehat{A'B'}N']$ et le trajet entre O et A ou A' d'une part, et entre B ou B' et N' d'autre part, étant unique et parfaitement déterminé, on a donc
$\mathrm{nch}(\widetilde{ON'}) = \mathrm{nch}(\widehat{AB}) + \mathrm{nch}(\widehat{A'B'}) = 2\,\mathrm{nch}(\widehat{AB})$ pour d'évidentes raisons de symétrie.
D'autre part, on a aussi, grâce à une tout aussi évidente invariance par translation du nombre de chemins entre deux points, $\mathrm{nch}(AB) = \mathrm{nch}(A'B') = \mathrm{nch}(ON)$.
On pose $\alpha = \mathrm{nch}(AB) = \mathrm{nch}(A'B') = \mathrm{nch}(ON)$.
De même soit $\beta = \mathrm{nch}(AB') = \mathrm{nch}(A'B)$ (ces deux nombres sont égaux par symétrie).
D'après le principe de symétrie, on a
$\mathrm{nch}(\widehat{AB}) = \mathrm{nch}(AB) - \mathrm{nch}(AB') = \alpha - \beta$, donc

$$\mathrm{nch}(\widetilde{ON'}) = 2\alpha - 2\beta. \tag{1}$$

Il se trouve qu'on peut aussi dénombrer les chemins joignant O et N' sans condition particulière en faisant intervenir α et β :
Il est clair qu'on a $[ON'] = [OABN'] \sqcup [OAB'N'] \sqcup [OA'BN'] \sqcup [OA'B'N']$, et donc

$\mathrm{nch}(ON') = \mathrm{nch}(AB) + \mathrm{nch}(AB') + \mathrm{nch}(A'B) + \mathrm{nch}(A'B') = 2\alpha + 2\beta$

$$\mathrm{nch}(ON') = 2\alpha + 2\beta \tag{2}$$

En ajoutant (1) et (2), on obtient
$\mathrm{nch}(ON') + \mathrm{nch}(\widetilde{ON'}) = 4\alpha = 4\,\mathrm{nch}(ON)$
Donc

$$\mathrm{nch}(\widetilde{ON'}) = 4\,\mathrm{nch}(ON) - \mathrm{nch}(ON') \tag{3}$$

□

V. Résolution des problèmes annoncés

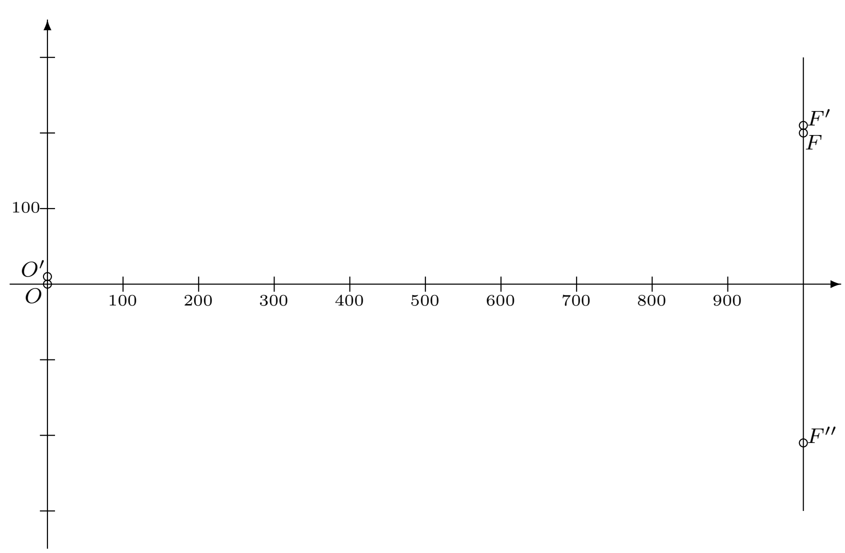
V.1. « Rappels » de dénombrement

On démontre en terminale que le nombre p parmi n vaut, pour $0 \leqslant p \leqslant n$ exactement

$$\binom{n}{p} = \frac{n\,!}{p\,!(n-p)!}$$

La notation $n!$ (factorielle de n) désigne le produit de tous les entiers entre 1 et n ; par exemple $4! = 1 \times 2 \times 3 \times 4$.
On a en particulier $n\,! = (n-1)! \times n$. On convient aussi que $1! = 0! = 1$.
D'autre part, le nombre de chemins de longueur q est exactement 2^q, puisque pour le premier « pas », il y a deux possibilités, puis pour le deuxième pas il y a encore deux possibilités (qui se multiplient : il y a 2×2 chemins de longueur 2), puis pour troisième pas il y a encore 2 possibilités et ainsi de suite.
Il y a donc $\overbrace{2 \times 2 \times \cdots \times 2}^{q \text{ facteurs}} = 2^q$ chemins de longueur q.

V.2. Solution du problème de dépouillement



Il y a autant de chemins joignant le point $O(0,0)$ et le point $F(1000,200)$ sans passer sous l'axe des abscisses que de chemins joignant le point $(O'(0,1))$ et le point $F'(1000,201)$ sans toucher l'axe des abscisses.
La probabilité cherchée est égale au quotient du nombre de chemins qui joignent O à F sans passer sous l'axe des abscisses par le nombre total de chemins joignant O et F :

$$P = \frac{\mathrm{nch}(\widehat{O'F'})}{\mathrm{nch}(O'F')} = \frac{\mathrm{nch}(O'F') - \mathrm{nch}(O'F'')}{\mathrm{nch}(O'F')} \qquad (F''(1000,-201) \text{ étant le point symétrique de } F').$$

La différence d'abscisse entre O' et F' ou entre O' et F'' est la même ($\Delta x = 1000$) ; la différence d'ordonnée entre O' et F' est $\Delta y = 200$, alors que la différence d'ordonnée entre O' et F'' est $\Delta y' = -202$. On a donc
$\mathrm{nch}(O'F') = \left(\begin{matrix}1000\\ \frac{1000+200}{2}\end{matrix}\right) = \binom{1000}{600}$
$\mathrm{nch}(O'F'') = \left(\begin{matrix}1000\\ \frac{1000-202}{2}\end{matrix}\right) = \binom{1000}{399}$
D'où

$$P = \frac{\binom{1000}{600} - \binom{1000}{399}}{\binom{1000}{600}} = \frac{\dfrac{1000!}{600!\,400!} - \dfrac{1000!}{399!\,601!}}{\dfrac{1000!}{600!\,400!}} = 1 - \frac{600!}{601!}\,\frac{400!}{399!} = 1 - \frac{400}{601} \simeq 0,33444$$

Comme $P \simeq \frac{1}{3}$, il y a donc environ une chance sur trois (seulement ?) pour que A ne soit jamais dépassé par B lors du dépouillement.

V.3. Solution du problème de fond de caisse

Ici encore, le nombre de chemins joignant A et B est simplement $\mathrm{nch}(AB) = \binom{200}{100}$, et pour déterminer le nombre de chemins dans $[AB]$ qui passent sous l'axe des abscisses, il suffit de surélever d'une case le problème pour pouvoir utiliser le principe de symétrie.
Soient $A'(0, n+1)$ et $B'(0, n+1)$; soit aussi $B''(0, -n-1)$ le symétrique de B' par rapport à l'axe des abscisses. Soit E_n l'événement « manque de monnaie » (lorsque le fond de caisse initial contient n billets).
En appliquant le raisonnement précédents, on trouve que
$$\mathrm{P}(E_n) = \frac{\mathrm{nch}(A'B'')}{\mathrm{nch}(A'B')} = \frac{\binom{200}{100+n+1}}{\binom{200}{100}} = \frac{100 \times 99 \times \cdots \times (100-n)}{(100+n+1)(100+n) \times \cdots \times 102 \times 101}$$

On ne peut pas résoudre de manière algébrique l'équation $\mathrm{P}(E_n) \leqslant 0{,}05$, mais par exemple par tâtonnement, on trouve que la première valeur de n qui convient est $n = 16$, pour laquelle on a $\mathrm{P}(E_{16}) \simeq 0{,}039$.

V.4. Solution du problème de marche aléatoire unidimensionnel

Nous sommes maintenant armés pour résoudre le problème de la marche aléatoire :
Modélisons tout d'abord un déplacement du marcheur aléatoire : on placera en abscisse le nombre de pas que celui-ci fait, et en ordonnée la distance algébrique qu'il a parcourue depuis sa maison, qui correspond donc à l'ordonnée 0. Un résultat de la marche aléatoire peut donc être modélisée exactement comme un chemin au sens des chemins étudiés dans la section précédente IV.

Désignons par $(R = 2n)$ l'événement « Le premier retour au point de départ a lieu après $2n$ déplacements ». L'événement « Le marcheur unidimensionnel aléatoire revient à son point de départ » se note R_0 et on a donc

$$R_0 = (R = 2) \sqcup (R = 4) \sqcup (R = 6) \sqcup \cdots \sqcup (R = 2n) \sqcup \cdots = \bigsqcup_{n=1}^{+\infty} (R = 2n).$$

R_0 est la limite, pour $n \to +\infty$ des événements emboîtés $T_n = (R \leqslant 2n)$, qu'on peut définir ainsi :

$$T_{2n} = (R = 2) \sqcup (R = 4) \sqcup (R = 6) \sqcup \cdots \sqcup (R = 2n) = \bigsqcup_{k=1}^{n} (R = 2k).$$

On admet que $\lim\limits_{n \to +\infty} \mathrm{P}(T_{2n}) = \mathrm{P}(R_0)$.
Soit S_k la position du marcheur après k déplacements.
$(S_{2n} = 0)$ désigne donc l'événement « Le marcheur est en 0 après $2n$ déplacements ».
En utilisant les notations qu'on vient de définir, on a donc $(S_{2n} = 0) = [ON']$, tandis que

$$\mathrm{P}(S_{2n} = 0) = \frac{\mathrm{nch}(ON')}{\text{nombre de chemins de longueur } 2n}$$

De même, $(R = 2n) = [\widetilde{ON'}]$ et $\mathrm{P}(R = 2n) = \dfrac{\mathrm{nch}(\widetilde{ON'})}{\text{nombre de chemins de longueur } 2n}$

Théorème 4 $\mathrm{P}(R = 2n) = \mathrm{P}(S_{2n-2} = 0) - \mathrm{P}(S_{2n} = 0)$

Démonstration : On considère que tous les chemins partant de O, de longueur $2n$ sont équiprobables. Soit K leur nombre. La probabilité d'un événement est égale au quotient par K du nombre de chemins différents qui réalisent cet événement.
L'événement $(S_{2n-2} = 0)$ est réalisé par les chemins qui passent par N. Mais un chemin joignant simplement O à N n'est pas de longueur $2n$, et il ne faut considérer que des chemins de longueur $2n$.
Il faut donc prolonger chaque chemin de $[ON]$, et il y a quatre façons de le prolonger, en repartant vers B ou vers B', puis en revenant vers N, ou en allant vers $C(2n, 2)$ ou vers $C'(2n, -2)$.
Donc le nombre de chemins de longueur $2n$ réalisant l'événement $(S_{2n-2} = 0)$ est égal à quatre fois le nombre de chemins joignant O à N.

Donc $\mathrm{P}(S_{2n-2} = 0) = \dfrac{4\,\mathrm{nch}(ON)}{K}$.

Plus simplement, on a $\mathrm{P}(S_{2n} = 0) = \dfrac{\mathrm{nch}(ON')}{K}$ et $\mathrm{P}(R = 2n) = \dfrac{\mathrm{nch}(\widetilde{ON'})}{K}$
En divisant par K la relation (3), on obtient exactement ce qu'on voulait démontrer :

$$\mathrm{P}(R = 2n) = \mathrm{P}(S_{2n-2} = 0) - \mathrm{P}(S_{2n} = 0).$$

Notons que pour appliquer cette formule pour $n = 1$, on aurait besoin de définir l'événement $(S_0 = 0)$. Nous pouvons admettre que cet événement est un événement certain, de probabilité 1.
Si cette possibilité nous gêne, on peut aussi le vérifier directement :
Il est clair que $(R = 2) = (S_2 = 0)$, et comme il y a 4 chemins de longueur 2, dont deux reviennent au point de départ, on a $\mathrm{P}(R = 2) = \mathrm{P}(S_2 = 0) = \frac{1}{2} = 1 - \mathrm{P}(S_2 = 0)$.

$\square$

V.4.1. Remarque

Le résultat que nous venons de démontrer permet d'établir que la suite des probabilités $\big(\mathrm{P}(S_{2n} = 0)\big)_{n \in \mathbb{N}}$ est une suite décroissante, ce qui n'a rien de trivial si on cherche à utiliser l'expression en fonction de n de chacun des termes.

V.4.2. Conséquence de cette relation

On a vu que l'événement « Le marcheur est repassé au moins une fois à la maison lorsqu'il en est au pas numéro $2n$ » (qu'on a noté T_{2n}) peut se décomposer ainsi :
$T_{2n} = R_2 \sqcup R_4 \sqcup R_6 \sqcup \cdots \sqcup R_{2n-2} \sqcup R_{2n}$.
Sa probabilité est donc

$$\begin{aligned}\mathrm{P}(T_{2n}) &= \big(1-\mathrm{P}(S_2=0)\big)+\big(\mathrm{P}(S_2=0)-\mathrm{P}(S_4=0)\big)+\big(\mathrm{P}(S_4=0)-\mathrm{P}(S_6=0)\big)+\dots\\ &\qquad\cdots+\big(\mathrm{P}(S_{2n-4}=0)-\mathrm{P}(S_{2n-2}=0)\big)+\big(\mathrm{P}(S_{2n-2}=0)-\mathrm{P}(S_{2n}=0)\big)\\ &= 1-\mathrm{P}(S_{2n}).\end{aligned}$$

Le nombre total de chemins de longueur $2n$ est $K = 2^{2n}$.
Le nombre de chemins entre les deux points accessibles de l'axe des abscisses $O(0,0)$ et $N'(2n,0)$ est, en application de la formule vue plus haut : $\binom{2n}{n}$.
Donc $\mathrm{P}(T_{2n}) = 1 - \dfrac{\binom{2n}{n}}{2^{2n}} = 1 - \dfrac{(2n)!}{2^{2n}(n\,!)^2}$.

On peut montrer, par exemple grâce à la formule de Stirling ($n! \sim \sqrt{2\pi n}\; n^n\, e^{-n}$) que la fraction ci-dessus tend vers 0, donc que $\mathrm{P}(T_{2n}) \to 1$. Je donne un peu plus loin une preuve « élémentaire » de ce résultat.
Ce résultat n'est pas franchement étonnant, on comprend bien que le nombre de chemins de longueur $2n$ joignant O et N' est très petit devant le nombre total de chemins de longueur $2n$, intuitivement, on a moins de chance de se retrouver en N' que de ne pas s'y retrouver.
On a donc $\lim\limits_{n\to+\infty} \mathrm{P}(T_{2n}) = 1$. Cela se traduit par le fait que la probabilité que le retour à la maison finisse par avoir lieu est 1. Au sens des probabilités, il est donc certain que le marcheur va finir par repasser au moins une fois devant sa maison, même si on ne peut pas savoir quand !

V.5. En guise de conclusion

On vient de démontrer que pour une marche unidimensionnelle, la probabilité d'un retour à la maison est 1.
On démontre qu'en dimension 2, ce résultat est encore vrai.
En revanche, en dimension 3, on démontre que la probabilité du retour à la maison n'est plus que de 0,3405…
D'où la célèbre phrase de Kakutani :
« *Un homme ivre pourra certainement retrouver le chemin de sa maison, alors qu'un oiseau ivre risque d'être perdu pour toujours* ».

VI. Annexe

VI.1. Un calcul de limite

Voici, grâce à une indication de Jean-Luc Hans, une démonstration « élémentaire » du fait que $\lim\limits_{n\to+\infty} \dfrac{\binom{2n}{n}}{2^{2n}} = \lim\limits_{n\to+\infty} \dfrac{(2n)!}{2^{2n}(n\,!)^2} = 0$.

Soit $(u_n)_{n\geqslant 1}$ la suite définie par $u_n = \dfrac{\sqrt{n}\binom{2n}{n}}{2^{2n}} = \dfrac{\sqrt{n}(2n)!}{2^{2n}(n\,!)^2}$.

Il est clair que (u_n) est une suite à termes positifs.

- Montrons tout d'abord que (u_n) est une suite croissante, en étudiant $\dfrac{u_{n+1}}{u_n}$

$$\frac{u_{n+1}}{u_n} = \frac{\dfrac{\sqrt{n+1}\big(2(n+1)\big)!}{2^{2(n+1)}((n+1)\,!)^2}}{\dfrac{\sqrt{n}(2n)!}{2^{2n}(n\,!)^2}} = \frac{\sqrt{n+1}(2n+2)!\,2^{2n}(n\,!)^2}{2^{2(n+1)}((n+1)\,!)^2\,\sqrt{n}(2n)!} = \frac{\sqrt{n+1}(2n+2)(2n+1)}{2^2(n+1)^2\sqrt{n}}$$
$$= \frac{2n+1}{2\sqrt{n(n+1)}} = \frac{n+\frac{1}{2}}{\sqrt{n^2+n}};$$

or il est clair que $(n+\frac{1}{2})^2 = n^2+n+\frac{1}{4} > n^2+n$, donc $n+\frac{1}{2} > \sqrt{n^2+n}$ et donc $\dfrac{u_{n+1}}{u_n} > 1$: la suite (u_n) est bien strictement croissante.

- Montrons maintenant par récurrence, que pour tout n, on a $u_n < \sqrt{\dfrac{n}{2n+1}}$;

Cette propriété est vraie pour $n = 1$: en effet, $u_1 = \dfrac{\sqrt{1}\binom{2}{1}}{2^2} = \dfrac{2}{4} = \dfrac{1}{2} = \sqrt{\dfrac{1}{4}} < \sqrt{\dfrac{1}{3}} = \sqrt{\dfrac{1}{2.1+1}}$.
Supposons maintenant que pour un certain $n \geqslant 1$, la propriété est vraie à l'ordre n, c'est-à-dire qu'on a $u_n < \sqrt{\dfrac{n}{2n+1}}$; montrons que la propriété est alors vraie à l'ordre $n+1$:
On a donc
$$u_{n+1} = \frac{u_{n+1}}{u_n}\,u_n = \frac{n+\frac{1}{2}}{\sqrt{n(n+1)}}\,u_n < \frac{2n+1}{2\sqrt{n(n+1)}}\sqrt{\frac{n}{2n+1}} = \sqrt{\frac{(2n+1)^2}{4n(n+1)}}\sqrt{\frac{n}{2n+1}}$$
Donc
$$u_{n+1} < \sqrt{\frac{2n+1}{4(n+1)}} = \sqrt{\frac{n+\frac{1}{2}}{2n+2}}.$$

On cherche à montrer que $u_{n+1} < \sqrt{\dfrac{n+1}{2(n+1)+1}} = \sqrt{\dfrac{n+1}{2n+3}}$; il suffit donc de prouver que $\dfrac{n+\frac{1}{2}}{2n+2} < \dfrac{n+1}{2n+3}$ pour achever la démonstration.
Or
$$\frac{n+\frac{1}{2}}{2n+2} - \frac{n+1}{2n+3} = \frac{(n+\frac{1}{2})(2n+3)-(n+1)(2n+2)}{(2n+2)(2n+3)} = \frac{(2n^2+4n+\frac{3}{2})-(2n^2+4n+2)}{(2n+2)(2n+3)}$$
$$= \frac{-\frac{1}{2}}{(2n+2)(2n+3)} < 0,$$
donc on a bien $\dfrac{n+\frac{1}{2}}{2n+2} < \dfrac{n+1}{2n+3}$ et $\sqrt{\dfrac{n+\frac{1}{2}}{2n+2}} < \sqrt{\dfrac{n+1}{2n+3}}$; puisqu'on avait $u_{n+1} < \sqrt{\dfrac{n+\frac{1}{2}}{2n+2}}$, on peut en déduire ce qu'on voulait démontrer : $u_{n+1} < \sqrt{\dfrac{n+1}{2n+3}}$.

On a donc bien prouvé par récurrence que pour tout $n \in \mathbb{N}$, on a $u_n < \sqrt{\dfrac{n}{2n+1}}$.

- Puisque $\sqrt{\dfrac{n}{2n+1}} < \sqrt{\dfrac{n}{2n}} = \sqrt{\dfrac{1}{2}}$, on a donc $u_n < \dfrac{1}{\sqrt{2}}$ pour tout n, donc la suite (u_n) est majorée; comme on a vu qu'elle était croissante, elle est donc convergente, et sa limite est

$\ell \leqslant \dfrac{1}{\sqrt{2}}$;

Puisque pour tout $n \geqslant 1$, on a $u_n = \dfrac{\sqrt{n}\binom{2n}{n}}{2^{2n}} < \dfrac{1}{\sqrt{2}}$, on en déduit que pour tout n, on a $0 < \dfrac{\binom{2n}{n}}{2^{2n}} < \dfrac{1}{\sqrt{2n}}$.

Puisque $\lim\limits_{n\to+\infty} \dfrac{1}{\sqrt{2n}} = 0$, par le théorème d'encadrement (ou « des gendarmes »), on peut conclure que $\lim\limits_{n\to+\infty} \dfrac{\binom{2n}{n}}{2^{2n}} = 0$. $\square$

VI.2. Un prolongement intéressant

Ce raisonnement sur le dénombrement des chemins, amène du coup une question dont je n'ai pas la réponse.
La conséquence de ce calcul est, puisque $\mathrm{P}(T_0 > 2n) = \mathrm{P}(S_{2n} = 0)$, qu'il y a autant de chemins de longueur $2n$ qui vont de O vers N' que de chemins de longueur $2n$ qui, partant de O, ne recoupent pas l'axe des abscisses. Il y a donc forcément une bijection entre l'ensemble $[ON']$ et l'ensemble des chemins de longueur $2n$ qui ne recoupent pas l'axe des abscisses. Je n'ai pas réussi à construire explicitement une telle bijection. Quelqu'un aurait-il une solution ?
J'ai réussi assez facilement à passer d'un chemin de longueur $2n$ ne rencontrant pas l'axe des abscisses à un chemin dans $[ON']$: il suffit de considérer, si le chemin se termine en $D(2n, 2p)$, la symétrie par rapport à la droite horizontale $y = p$, et d'appliquer cette symétrie à une partie du chemin comprise entre D et un des points d'intersection de cette droite avec le chemin. Mais quel point d'intersection choisir ? Le premier ? Le dernier ? Et si on est sûr d'obtenir ainsi, quel que soit le choix qu'on fasse, un chemin qui arrive à N', comment être sûr du caractère injectif et/ou surjectif d'une telle application ?
La résolution de ce nouveau problème, la construction effective d'une bijection adéquate nous donnerait forcément une nouvelle solution « brillante » de ce problème de marches aléatoires !