La marche au hasard

« C'est curieux chez les marins ce besoin de faire des phrases… »

Francis Blanche, Les Tontons flingueurs (1963), écrit par Michel Audiard

Les marins, dans les ports, sont de façon traditionnelle sujets à un petit jeu appelé souvent en sciences « la marche du marin ivre ». Bien sûr on pourrait étendre cette occupation nocturne à d'autres parties de la population comme les étudiants ou les supporters, mais qu'importe. Conservons l'image familière du marin totalement ivre devant rentrer à bord en fin de nuit après avoir ingurgité une ration tout à fait excessive d'alcool…

Précisons alors que notre marin ne maîtrise plus vraiment ses fonctions motrices et déambule totalement au hasard d'un pas vers le suivant. Supposons plus précisément qu'à chacun de ses pas il se dirige de façon totalement aléatoire et équiprobable vers le Nord, le Sud, l'Est ou l'Ouest comme le représente la *figure 4.1*.

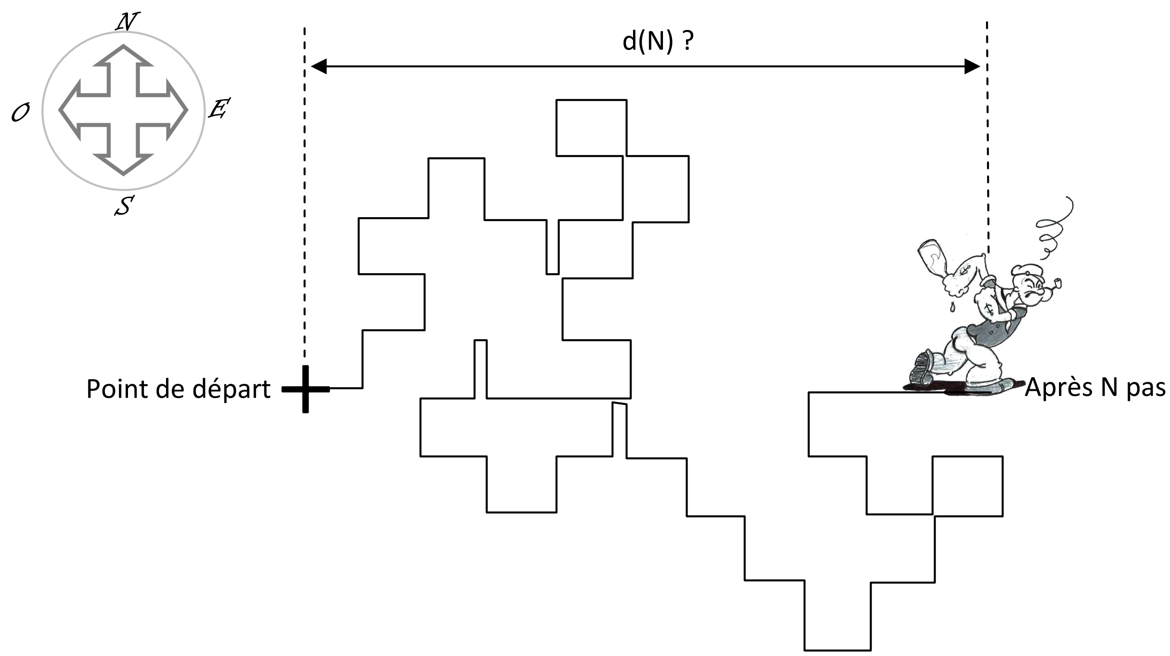


Figure 4.1 : La marche compliquée du marin ivre durant 60 pas

La question qui vous est alors posée est la suivante :

« Pouvez-vous déterminer, au bout d'un certain nombre N de pas, où se trouvera statistiquement le marin par rapport à son point de départ ? »

Tout d’abord, posons correctement le problème :

Les pas du marin sont supposés de même amplitude, et dirigés de façon totalement aléatoire et équiprobable vers un des 4 points cardinaux. On s’intéressera, pour un nombre N entier de pas, à déterminer la distance d(N) à laquelle devrait être le marin.

- En priorité il faut s’approprier le problème… en commençant avec un cas simple. Déterminez alors, pour un mouvement uniquement vertical sur l’axe Nord/Sud, quelle grandeur vous permettrait, au bout de N pas d’appréhender le déplacement. Pensez à la moyenne du déplacement, à la moyenne de son carré, etc… Essayez d’établir un lien avec la probabilité de chaque déplacement.
- Une fois la grandeur caractéristique et sa loi de probabilité déterminée, étendez le problème aux quatre directions proposées.
- Précisez ainsi, de façon la plus claire possible, s’il est possible ou pas d’avoir une idée de la distance d(N) pour une course au hasard particulière, ou pour un ensemble de courses au hasard.

Saurez-vous alors préciser s’il est possible de prévoir la distance à laquelle sera le marin après N pas ?

Solutions :

Marche au hasard en une dimension (1D) :

Commençons par envisager un déplacement au hasard, mais uniquement dans la direction verticale, dans l'axe Nord / Sud. A chaque pas, notre marin va vers le haut ou vers le bas, et le nombre de possibilités correspondant aux premiers pas est alors facile à identifier et va nous aider à trouver quelle grandeur est caractéristique.

La *figure 4.2* illustre ce qui peut se produire le long des 5 premiers pas en présentant en ordonnée les déplacements verticaux possibles et en abscisse l'indice des pas.

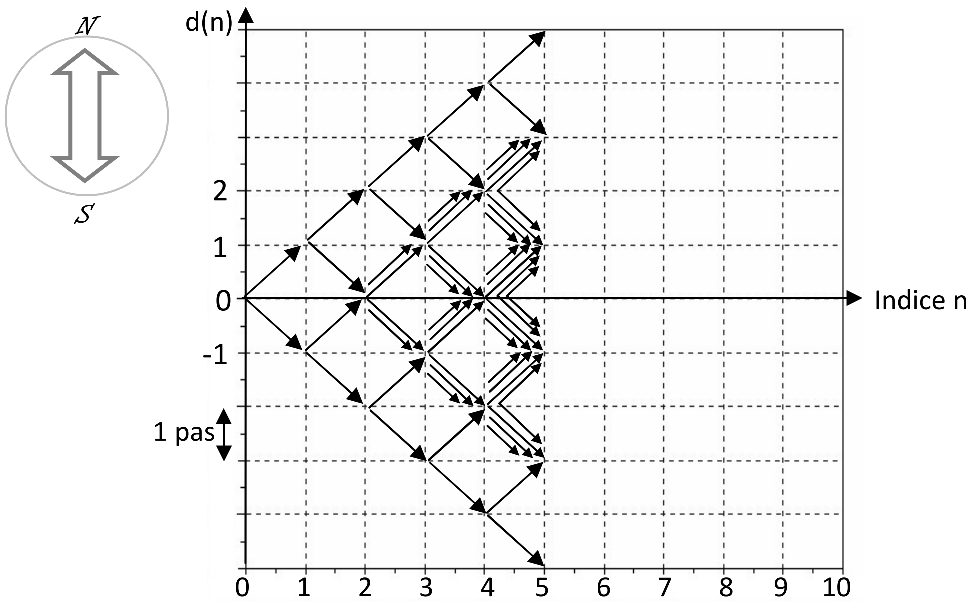


Figure 4.2 : Marche au hasard verticale sur 5 pas

Après le pas n°1 : le déplacement d_1 peut prendre la valeur +1 ou la valeur -1. La moyenne statistique du déplacement, autrement dit son « espérance mathématique » s'écrit :

$$E(d_1) = \frac{(+1)+(-1)}{2} = 0.$$

Cette grandeur n'a donc pas l'air d'être très intéressante puisqu'elle s'avère également nulle pour toutes les autres valeurs du nombre de pas (la symétrie du graphe par rapport à l'axe le montre à l'évidence).

Dans ce cas de figure, et de façon très répandue dans les domaines scientifiques et dans les calculs statistiques, il convient alors de s'intéresser à la moyenne <u>du carré</u> des valeurs, ce qui permet de s'affranchir du signe des déplacements et de s'intéresser ainsi à la « distance quadratique moyenne ». La grandeur appréhendée est alors appelée la « variance ».

Pour le premier pas, elle s'écrit : $E\left({d_1}^2\right) = \frac{1^2+(-1)^2}{2} = 1$

Après le pas n°2 : $E\left({d_2}^2\right) = \frac{2^2+0^2+0^2+(-2)^2}{4} = 2$

Après le pas n°3 : $E\left({d_3}^2\right) = \frac{3^2+3\times1^2+3\times(-1)^2+(-3)^2}{8} = 3$

NB : il faut bien comprendre ce qui se passe à partir du 3ème pas : il y avait 2 possibilités pour atteindre la valeur 0 au pas d'avant, chacune d'elle donne une possibilité d'atteindre le +1 au pas n°3. En ajoutant la possibilité du déplacement de +2 jusqu'à +1, cela fait en tout 3 possibilités pour obtenir la valeur +1 au pas n°3. Idem pour la valeur -1. Il ne faut donc pas juste compter les valeurs atteignables, mais toutes les possibilités permettant de les atteindre. C'est cela qu'il faut ensuite moyenner.

Ainsi, en généralisant, et en vérifiant que ca fonctionne toujours pour les indices supérieurs sur la *figure 4.2*, on obtient la relation :

$$E\left({d_n}^2\right) = n$$

En d'autres termes, la variance de la distance atteinte pour n pas, sur un grand nombre d'essais, tendra vers *n* le nombre de pas réalisé.

NB : Il est très répandu d'évoquer « l'écart type » de la variable, qui est simplement la racine carrée de la variance : $\sigma = \sqrt{E\left({d_n}^2\right)} = \sqrt{n}$.

NB : Une précision également : la variance, de façon générique, s'écrit comme la moyenne des carrés des écarts à la moyenne : $V_n = E((d_n - E(d_n))^2)$. *Dans nos calculs précédents, la moyenne étant nulle, cette grandeur revient alors à :* $E\left({d_n}^2\right)$.

Il faut bien comprendre à ce stade que ce qui est prévisible, c'est uniquement que sur un grand nombre d'essais la variance tende vers le nombre de pas réalisé dans l'essai. Il n'est pas contre pas du tout possible de prévoir la distance ou même son carré pour un essai particulier !

Autrement dit, la réponse est déjà trouvée : Non il n'est pas possible de prévoir la distance à laquelle sera le marin après N pas. Par contre, en faisant l'expérience sur un grand nombre de marins, et en mesurant la distance de chacun après N pas, et en moyennant les carrés de ces valeurs, on obtiendrait à peu près N.

Il est d'ailleurs possible de simuler ce petit problème de manière à valider le raisonnement et la formule obtenue. Un petit programme réalisé sous le logiciel Scilab® permet de tracer les évolutions des distances pour plusieurs marins et pour un certain nombre de pas. La *figure 4.3* représente une simulation particulière effectuée pour 10 marins faisant chacun 30 pas.

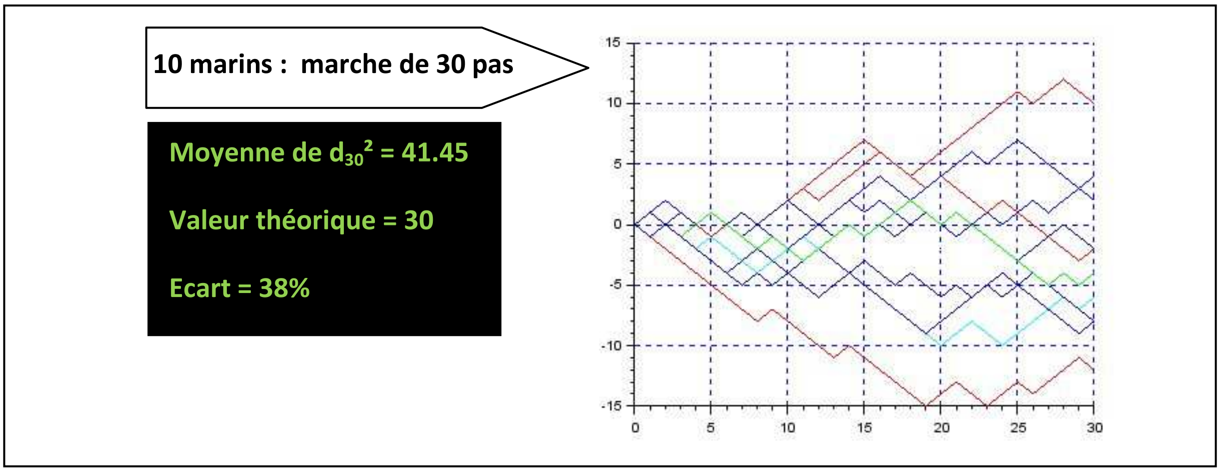


Figure 4.3 : Simulation pour 10 marins sur 30 pas

Avec seulement 10 marins pour faire le test, la valeur de la moyenne des carrés des distances simulées s'écarte assez fortement de la valeur théorique qui vaut 30.

Afin d'observer la convergence de la variance, il est nécessaire d'effectuer un grand nombre de tests. La *figure 4.4* représente alors les trajectoires obtenues pour 1000 marins faisant 30 pas, puis pour 10000 marins faisant 100 pas.

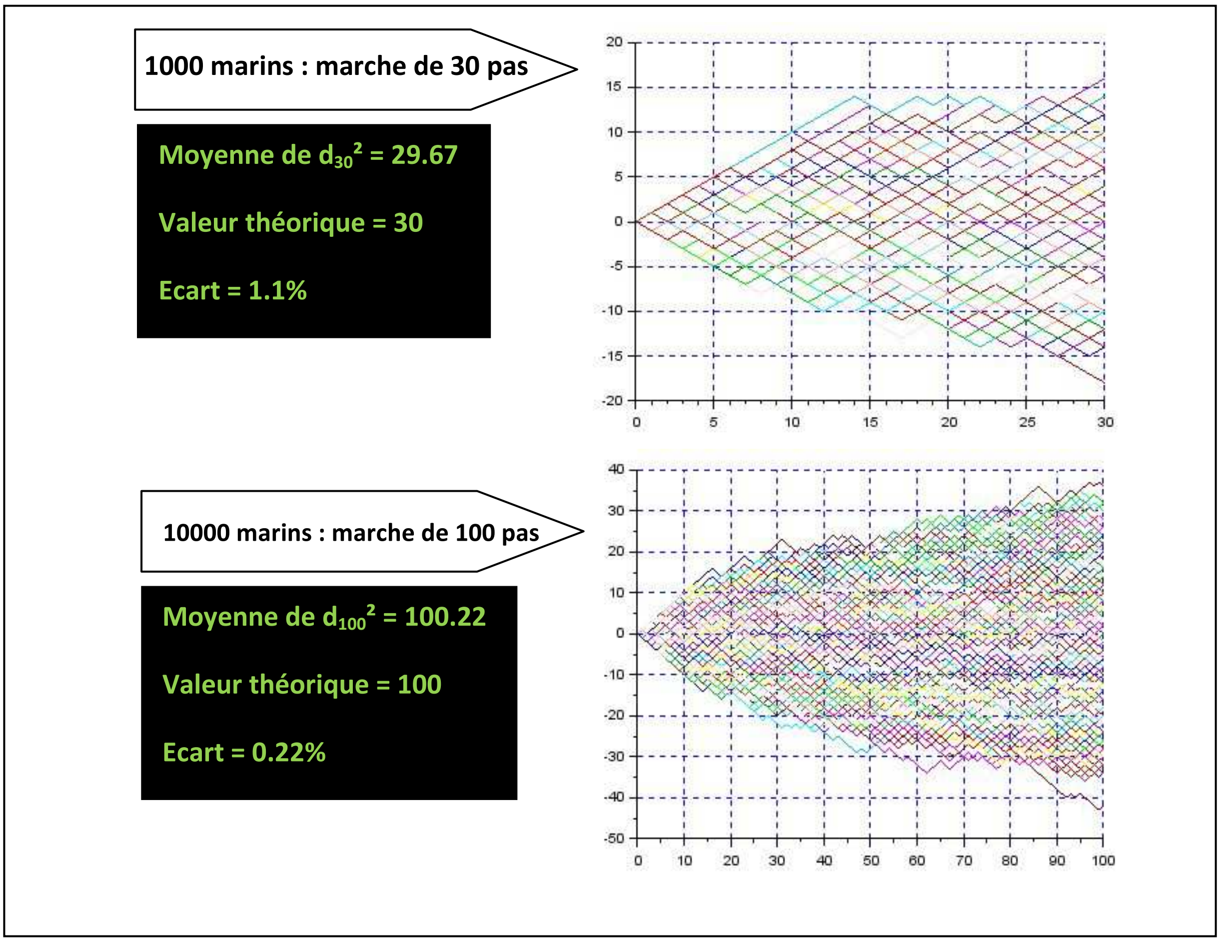


Figure 4.4 : Simulation pour 1000 marins sur 30 pas puis 10000 marins sur 100 pas

Les valeurs calculées et les valeurs théoriques des variances correspondent alors d'autant plus étroitement que le nombre de marins (le nombre d'essais différents) est important, ce qui est bien conforme au principe dit « des grands nombres » souverain en statistiques.

Marche au hasard en deux dimensions (2D):

Même si la réponse à la question posée est déjà éludée, il est intéressant de réfléchir à la marche au hasard en deux dimensions évoquée de façon initiale dans le problème. Dans ce cas là, il suffit de considérer qu'à chaque pas le marin va vers le Nord, ou vers le Sud, ou vers l'Est, ou vers l'Ouest. Son déplacement ressemblera alors au trajet suivi sur la *figure 4.5*.

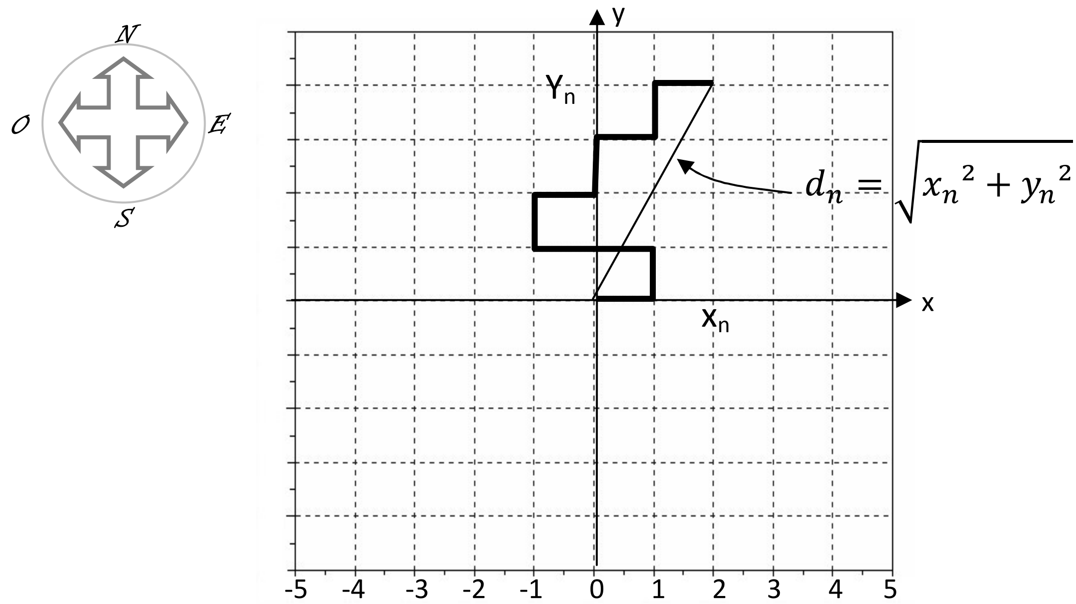


Figure 4.5 : Marche au hasard en 2D sur 10 pas

La figure illustre le fait que la distance effectuée pour le pas n vérifie : ${d_n}^2 = {x_n}^2 + {y_n}^2$.

Ainsi, on peut écrire : $E({d_n}^2) = E({x_n}^2 + {y_n}^2) = E({x_n}^2) + E({y_n}^2)$

En adoptant pour chaque coordonnée l'identification des variances à la valeur des indices, il vient : $E({d_n}^2) = 2.n$

Il ne reste plus qu'à confirmer la validité de cette égalité en effectuant des simulations portant sur un grand nombre d'essais. La *figure 4.6* représente ainsi la superposition de 1000 trajets au hasard réalisés à partir du point (0,0) et portant chacun sur 100 pas.

Le calcul de la variance restitue un résultat tout à fait conforme à celui prévu, avec une marge d'erreur comparable à celle déplorée sur la *figure 4.4* pour 1000 essais (environ 1%).

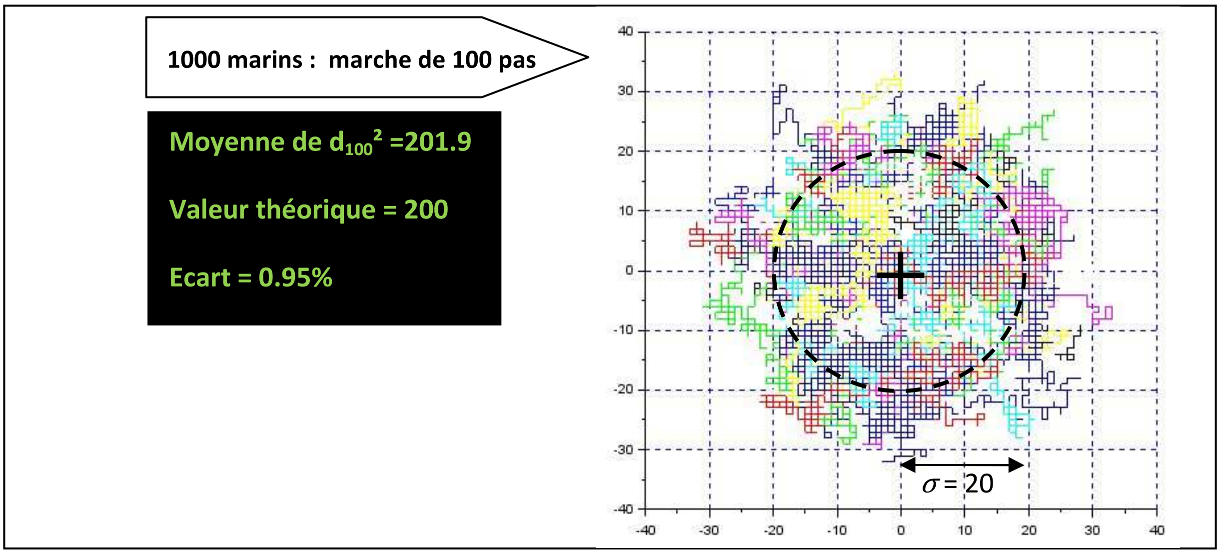


Figure 4.6 : Simulation 2D pour 1000 marins sur 100 pas

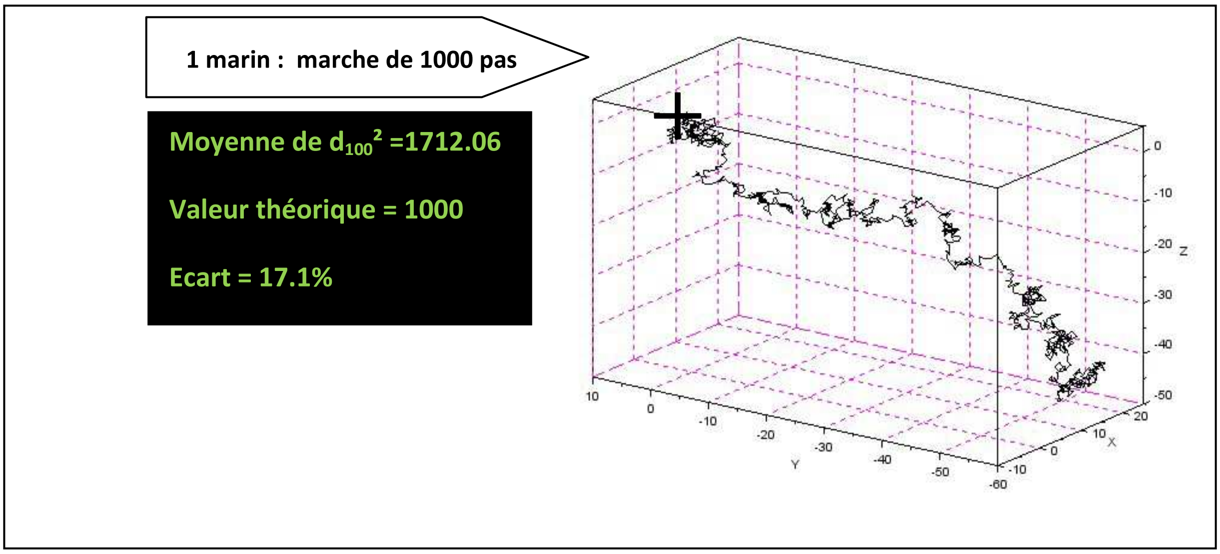
En calculant l'écart type dans ce dernier cas on obtient : $\sigma = \sqrt{E({d_n}^2)} = \sqrt{2.n} = 20.$

Cette valeur précise, homogène à une distance, permet de tracer un cercle de rayon σ qui situe statistiquement la majorité des trajets. Il est important alors de bien comprendre que si on ne sait pas prévoir le déplacement d'un marin donné, on connait tout de même avec une assez bonne précision le lieu des trajectoires probables grâce au calcul de la variance et de l'écart type.

Marche au hasard en trois dimensions (3D):

Pour finir, il est tout à fait possible d'étendre les raisonnements précédents à un déplacement en 3 dimensions. La variance après *n* pas serait alors de l'ordre de $E({d_n}^2) = 3.n$ et même en supposant des modes de déplacement aléatoires différents, on retrouverait le fait que l'écart type du déplacement autour de la moyenne varie « en $\sqrt{n}$ ».

C'est le cas de la simulation de la *figure 4.7* qui présente le cas d'une seule trajectoire aléatoire de 1000 pas, puis la superposition de 1000 trajectoires de 1000 pas chacune.



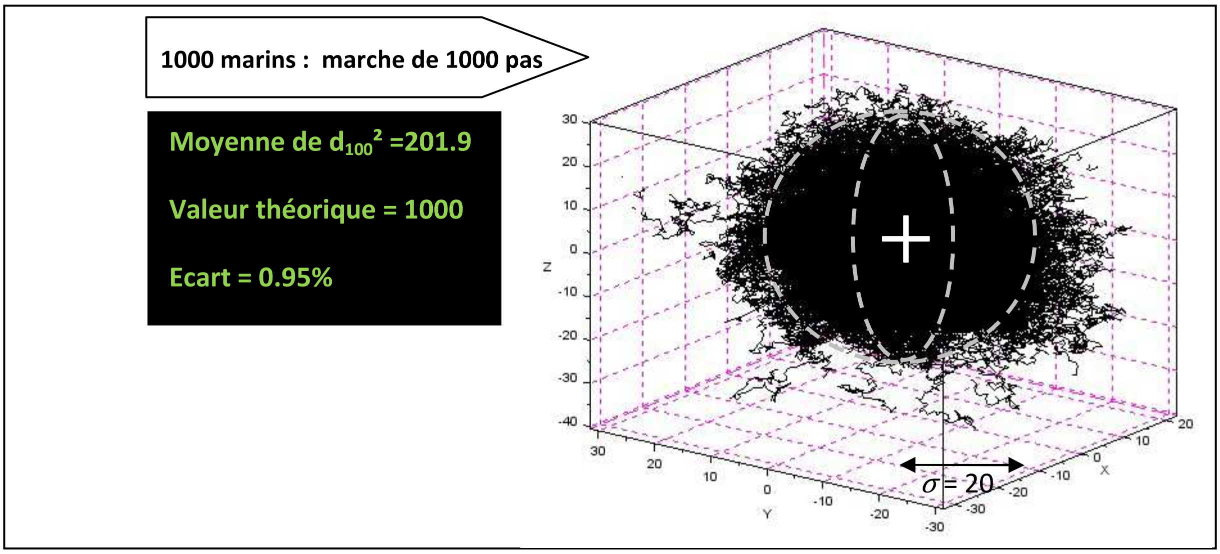


Figure 4.7 : Simulation 3D pour 1000 pas 1 trajectoire puis 1000 trajectoires

Cette simulation utilise un déplacement constitué, à chaque pas, d'un saut aléatoire sur chaque coordonnée (c'est plus rapide à exécuter pour un grand nombre d'itération que la stratégie précédente). On montre alors que se variance est : $E({d_n}^2) = \mathrm{n}$. En calculant l'écart type dans ce dernier cas on obtient : $\sigma = \sqrt{E({d_n}^2)} = \sqrt{n} = 31{,}62$, d'où le rayon délimité sur la figure qu'il faut envisager cette fois comme le rayon d'une « sphère de probabilité des trajectoires ».

Ce graphe fait vraiment penser à l'agitation chaotique de particules qui, individuellement sont imprévisibles, mais occupent un volume finalement délimité de façon claire par le rayon de la sphère qui évolue en $\sqrt{n}$.

Pour la petite histoire …

L'étude de la marche aléatoire a fait l'objet de beaucoup de développements et d'applications. La plus connue est le « mouvement brownien », ou encore le mouvement désordonné de particules dans un fluide isotrope détecté par un biologiste du nom de Richard Brown. L'étude de ce mouvement a mis en lumière le fait que sa variance soit proportionnelle au temps écoulé (c'est évident si la durée des pas est une constante) et les différents aspects statistiques et probabilistes de tels mouvements ont inspiré Einstein pour déterminer une formule donnant le nombre d'Avogadro, pour exprimer par des expériences la réalité des atomes et plus tard ont conduit Norbert Wiener à développer la stochastique et l'étude du bruit, etc… Faites quelques recherches et vous verrez tout ce que cette marche du marin ivre a induit en sciences.