

Théorie du petit trou et ses applications en biologie

J. Cartailier^{*}, J. Reingruber^{*} D. Holcman^{*,+}

^{*}Groupe de Math. Appliquées et Biologie Computationnelle, ENS Paris,
France

MANUSCRIPT CONFIDENTIEL, Pour DIFFUSION RESTREINTE-évaluation

Résumé :

Cet article décrit la théorie mathématique du petit trou (Narrow Escape theory [1]) et ces applications en biologie. Cette théorie, introduite il y a 15 ans [2], a pour but de trouver un cadre de calcul pour dériver les formules mathématiques qui estiment le temps pour qu'une petite particule (molécule, ion, protéine) qui bouge de façon aléatoire, trouve une petite porte de sortie dans un domaine confiné. Ces formules sont obtenues grâce à la théorie asymptotique des équations aux dérivées partielles. Elles permettent de relier la dynamique aléatoire du mouvement à la géométrie du domaine. Ces formules sont cruciales et difficile à explorer par d'autres méthodes (numériques ou expérimentales) car le temps de sortie tend vers l'infini quand la taille de la fenêtre devient toute petite. Ces formules révèlent par exemple comment la géométrie régule les flux en biologie, comment les réactions chimiques stochastiques ont lieu en milieu confiné, comment la vitesse de transduction du signal dépend du nombre de molécules... Ces formules nous aident donc à mieux comprendre et quantifier des phénomènes du vivant ayant lieu dans des compartiments biologiques de la taille allant du nano au micromètre, qui restent encore difficilement accessibles à l'expérimentation. Au fond, cette approche mêlant modélisation, analyse mathématique, simulations de l'aléatoire, analyse de grands jeux de données permet de replacer les mathématiques appliquées sur la scène de la science prédictive et de la découverte en biologie et non simplement descriptive.

Introduction

L'évolution de la vie sur terre a depuis plus de quatre milliards d'années créé des formes géométriques surprenantes (Fig. 1) et sans doute optimales pour l'adéquation de l'organisme à son environnement, et ce, du nanomètre, en passant par des formes complexes cellulaires, jusqu'aux grands mammifères d'aujourd'hui. Mais quelle est la fonction de ces formes géométriques au-dessous de l'échelle de la cellule vivante ? Comment ces formes permettent-elles aux cellules de fonctionner correctement ou même de façon optimale ? Si le domaine de la géométrie mathématique s'intéresse à classifier les formes en général, en les regroupant en fonction de leur topologie, des distances intrinsèques, etc... La physiologie, elle, a pour but de comprendre la fonction de la cellule. Mais alors comment relier la forme ou la morphologie cellulaire à la fonction biologique ?

La nouvelle biophysique mathématique tente de répondre à ces questions en formulant des lois cellulaires à partir de la physique statistique moléculaire élémentaire. Cet article présente une étape importante de cette approche basée sur la théorie du petit trou. Cette théorie consiste à calculer le temps que met une particule, molécule ou un objet bougeant de façon aléatoire dans un domaine clôturé, à trouver une petite porte de sortie. Obtenir une formule mathématique qui relie les paramètres géométriques du domaine dans lequel évolue la particule (comme le volume, son diamètre, les régions de rétrécissement à petit diamètre, les couloirs, etc...) mais aussi les caractéristiques de la porte de sortie, permettent d'établir les lois biophysiques des cellules, pour la régulation des flux moléculaires, des réactions chimiques et les messages entre cellules ou différents sous-compartiments. En particulier, nous verrons aussi qu'il existe des couloirs de rétrécissement, dit couloirs du temps qui font perdre un temps quasiment infini. Le rôle de ces couloirs est équivalent, pour le calcul du temps moyen, à réinitialiser la position de la particule là où elle a commencé une infinité de fois.

Dans la première partie de cet article, nous aborderons les aspects de la formulation probabiliste de la théorie du petit trou, puis nous discuterons la relation entre cette théorie et les événements rares. Ensuite nous décrirons les méthodes mathématiques de résolution. Enfin dans les dernières parties, nous donnerons plusieurs exemples d'application en neurobiologie, comme la régulation du calcium dans les synapses, la biophysique et l'électricité et enfin nous aborderons le rôle des comportements extrêmes dans la signalisation biochimique et le processus de recherche d'un ovule dans la fertilité.

S'il n'y avait qu'une chose à retenir, nous dirions que des nouveaux progrès en biologie et en médecine vont pouvoir être obtenus par le développement de théories, de calculs et de développements algorithmiques, de simulations et d'analyses des données, et aussi (*trouver*) par la découverte des formules qui révèlent le lien intime entre la géométrie et la fonction biologique. Il n'y a pas de beauté du vivant, il n'y a que des formes optimales, dite parfaites, qui remplissent efficacement leur fonction.

1-Qu'est-ce que la théorie du petit trou ou l'histoire de l'échappée belle aléatoire ?

La première question est d'ordre mathématique [1-2]: combien de temps faut-il à une particule infiniment petite et qui bouge de façon aléatoire à sortir par un petit trou circulaire de rayon petit (que nous appelons epsilon) ? Par exemple, une mouche qui entre dans votre appartement, quel temps va-t-elle mettre à sortir par une fenêtre ouverte ? et, que se passe-t-il quand l'ouverture de la fenêtre devient toute petite ? ... Et s'il y a un courant d'air, etc. Dans la limite où le trou devient

tout petit, le temps moyen de la mouche tend vers l'infini, mais de quelle manière ? ... En fonction de la taille epsilon, du reste du lieu et de son mouvement ?

Cette question, *qui est* d'ordre probabiliste, peut être résolue d'abord en la formulant comme une équation aux dérivées partielles, puis en utilisant des méthodes asymptotiques qui sont des outils mathématiques effectifs pour trouver des solutions de ces équations, quand un petit paramètre (ici l'ouverture de la fenêtre) tend vers zéro. Il est important de trouver des solutions explicites si on veut ensuite trouver une formule. Dans la boîte 1, nous avons résumé les équations et leur résolution.

2- La physique statistique des événements rares et le lien avec la théorie du petit trou

La recherche d'une petite porte de sortie par un processus aléatoire [4-5-10], qui n'a pas la faculté de voir cette porte mais de la reconnaître quand il est dessus, est intimement liée à la notion d'événements rares et importants. En effet, trouver la sortie prend beaucoup de temps et c'est aussi l'événement majeur de cette recherche.

Qu'est-ce qu'un événement rare ? C'est un événement qui n'arrive qu'une certaine fraction de fois d'événements élémentaires, disons une fois tous les 100, 1000 ou 10^9 . Par exemple, parmi toutes les collisions possibles, une collision sur 1000 ou 10000, en fonction de la concentration d'une molécule, va donner une réaction chimique. Cet exemple est générique en physique statistique. Quand un avion est soumis à beaucoup de vibrations, combien de temps faut-il attendre avant qu'une pièce ne se fracture ? Un autre exemple : parmi toutes les rencontres possibles pour un individu, seulement une petite fraction pourra donner naissance à une relation de longue durée. Dans tous ces cas, ce qui est intéressant est de savoir combien de temps il faudra attendre pour voir apparaître ces événements rares. Pour calculer ces temps d'attente, il faut alors soit disposer d'un modèle pour les prédire, *ou* soit construire une distribution empirique *quand* on peut mesurer directement la distribution de ces temps. Cette approche est une partie importante de la physique statistique qui a pris un essor avec les premiers travaux de Smoluchowski 1916, Kramer 1940, Chandrasekhar 1943 et Matkowski-Schuss 70s.

Trouver une petite sortie, une cible, peut aussi se formuler comme la première visite d'une petite fraction de l'espace (ou encore dit espace des phases). Le temps de la première visite est exponentiellement distribué, caractérisée par sa moyenne. Le calcul de ce temps a été possible en utilisant la théorie mathématique asymptotique, résumée dans la boîte 1 et 2 (classification des formules).

3- Transformation d'une question de probabilité en une question d'analyse mathématique et de résolution asymptotique

Après avoir formulé la théorie du petit trou comme la recherche d'une cible par un mouvement aléatoire, une seconde étape consiste à retransformer cette question de probabilité en une question d'analyse mathématique. Ce qui est possible par la théorie du calcul d'Ito développé dans les années 50. Dans ce contexte d'analyse mathématique, cette question revient à résoudre une équation aux dérivées partielles (boîte 1). Pour cela, il a fallu adapter des méthodes effectives de résolution basée sur la théorie des couches limites utilisée pour étudier le mouvement des fluides proches du bord d'une rive. Cette approche a surtout été développée par les écoles de mathématiques anglaises : Trois méthodes furent développées : l'une basée sur les fonctions de Green, qui est utilisée en seconde année d'université pour calculer le voltage d'une seule charge électrique placée dans le vide. Cette méthode permet de dériver des formules qui

montrent comment le temps de recherche diverge vers l'infini quand la taille de la fenêtre tend vers zéro. Ce temps dépend aussi de la géométrie locale, mais aussi, s'il y en a, de longs couloirs, appelés couloirs du temps et qui ralentissent grandement la recherche. Les deux autres méthodes s'appellent respectivement l'une, la méthode de Collins développée dans le paragraphe [2], l'autre *est* la méthode d'assemblage asymptotique développée par J. Keller (2017) et M. Ward à l'Université de Vancouver. Nous renvoyons le lecteur intéressé aux boîtes 1 et 2 pour la formulation et la classification des formules.

4- Application de la théorie du petit trou au trafic des molécules, à la modélisation et au calcul du temps de résidence du calcium dans les synapses.

Nous avons vu que les événements sont dits rares quand ils n'apparaissent qu'avec une échelle de temps bien plus longue que ceux décrits par l'ensemble des événements qui sont considérés. Mais si les événements deviennent trop rares devant la fonction qui doit être accomplie, alors celle-ci ne peut pas avoir lieu. Autrement dit, les échelles de temps doivent être grandes mais pas trop grandes. Alors qu'est-ce qui définit ces temps longs mais pas trop longs et comment les reconnaît-t-on en biologie cellulaire ? Voici quelques exemples :

Combien de temps *restent* les ions calcium dans une synapse après *qu'ils y soient* entrés (Figs. 2 et 3) ? Combien de temps une molécule Brownienne confinée dans un domaine membraneux met-elle à sortir ? Combien de temps met une protéine juste après avoir été synthétisée à trouver sa place dans la cellule ? Combien de temps les premières protéines, dites G, mettent-elles une fois produites à activer les récepteurs localisés un peu plus loin de la source de production dans la cellule [7] ? Ces temps-là ne sont pas calculés par les théories de la réaction chimique classique, ni par la théorie de la réaction-diffusion, mais par la statistique des premiers temps, dans la théorie du petit trou.

Le temps transitoire de résidence du calcium dans les synapses définit la possibilité de changer la géométrie d'une partie de la synapse et donc la force de modulation du neurone par la synapse, ce qui est la base de l'apprentissage et de la mémoire au niveau neuronale. Ces temps sont résumés dans le Boîte 2 et nous allons maintenant les commenter.

Voici quelques exemples :

Si le temps moyen que passe un récepteur dans la synapse définit la puissance du poids synaptique, alors il est important d'estimer ce temps. Pour cela il faut un modèle pour calculer ce temps en dimension deux : une première idée était de construire un domaine Ω de confinement entouré d'une barrière sauf à un endroit où il y a une porte de sortie. Il est aussi possible de mettre plusieurs portes [1-4] (Figs. 2-3). Nous sommes là par excellence dans la situation du petit trou. Alors le temps moyen que met une particule Brownienne qui commence au centre d'un disque à sortir par une petite fenêtre est donnée par la formule :

$$E[\tau | x(0) = \text{Centre}] = \frac{|\Omega|}{\pi D} \left[\log \frac{1}{\varepsilon} + \log 2 + \frac{1}{4} + O(\varepsilon) \right], \quad (1)$$

ou $|\Omega|$ est la surface du domaine, D le coefficient de diffusion, ε le ratio de la longueur du périmètre réfléchissant sur la longueur de la fenêtre absorbante. La formule est valable quand ε est petit. Dans le cas d'un domaine général, le premier terme reste valable mais le second change et dépend de la géométrie du domaine. Il est possible de pousser les calculs plus loin [3] et de voir alors comment le temps dépend de la position initiale de la molécule, mais ce terme contribue très peu au temps passe dans le domaine.

Un autre exemple cette fois en trois dimension est de comprendre qu'est ce qui définit le temps moyen des ions calcium dans une synapse ? Pour cela nous devons trouver une autre formule pour une structure qui contient maintenant un long cou (Fig. 3) qui peut se connecter soit de façon lisse soit de façon abute, alors dans ce cas, une trajectoire qui sort par le cou, peut malheureusement revenir une infinité de fois dans la tête, ce qui peut rajoute beaucoup de temps, du coup, le temps passe de

$$\tau_{3\text{dim}} = \frac{|\Omega|}{4aD} \left[1 + \frac{a}{\pi R} \log \frac{R}{a} + O(1) \right],$$

Dans le cas d'une boule de rayon R, de volume $|\Omega|$ en dimension 3, a est le rayon d'une petite fenêtre circulaire et D le coefficient diffusion. Maintenant quand un coup de longueur se rajoute alors la formule devient [1]

$$\tau_{3\text{dim}} = \begin{cases} \frac{|\Omega_1|L}{\pi a^2 D} + \frac{|\Omega_1|}{4aD} \left[1 + \frac{a}{\pi R} \log \frac{R}{a} \right] + \frac{L^2}{2D} & \text{pour une boule avec un cou} \\ \frac{|\Omega_1|L}{\pi a^2 D} + \frac{1}{\sqrt{2}} \left(\frac{R_c}{a} \right)^{3/2} \frac{|\Omega_1|}{R_c D} + \frac{L^2}{2D} & \text{pour un domaine quelconque (liaison avec un cusp)} \end{cases}$$

A note le terme en $1/a^2$, qui croit beaucoup plus vite que $1/a$ quand a tend vers zéro. Ces formules en combinaison avec des expériences ont permis de montre que la géométrie des synapses peut donc moduler le temps que passe les ions calcium dans la partie postérieure d'une synapse. Nous savons que ce temps est crucial pour l'induction de changement du poids synaptique.

Evaluer le temps d'agrégation d'un récepteur à la surface des cellules immunitaires permet aussi de définir la stratégie de recherche d'une cellule étrangère et de les différencier *avec les* cellules du corps, tout en scannant le maximum possible de cellules. Enfin, quand ces temps sont modifiés par les changements de la forme, alors la fonction change. Comprendre les mécanismes qui maitrisent et contrôlent ces temps nous permet aussi de prédire la disposition fine des structures et agencement moléculaire qui peuvent être ensuite vérifiées par microscopie [8]. Autrement dit, cette approche comparative entre l'analyse de formule, leur simulation a une véritable force d'explication et prédictive pour découvrir l'organisation fine du vivant au niveau moléculaire.

4-L'organisation du noyau d'une cellule, le temps pour qu'un facteur de transcription TF trouve ça cible et les modelés de polymère, comme Avatar de la chromatine dans les simulations informatiques.

Le noyau des cellule Eucaryote contient le matériel génétique (Fig. 4a-b), l'ADN qui s'organise comme de long spaghetti tous mélangés dans une grande sauteuse en cuivre bouillonnante, avec néanmoins certaines différences. Notamment la chromatine qui représente l'ADN avec des molécules contiennent des régions plus concentrées, avec plus de boucles. Une question importante est de savoir combien il y a de boucles, leur taille. Y a-t-il plus de boucle la ou les gènes sont exprimés, et combien faut-il de temps pour former une boucle? Le temps pour former une boucle correspond a trouvé un petit trou dans un espace abstrait : temps pour que les deux extrémités se touche pour la première fois [11].

De façon intéressante, la réponse d'une cellule à un stimulus externe nécessite souvent l'activation ou la répression de gènes spécifiques. Ce processus est régulé par des protéines appelées facteurs de transcription (TF) qui doivent trouver et se lier au promoteur de l'ADN de ces gènes (Fig. 4c-d), activant ou réprimant ainsi la transcription des gènes. Le défi pour ces TFs consiste à trouver le petit site promoteur composé d'environ 20 paires de bases d'ADN dans un génome de plusieurs millions de paires de bases. Comment cela est-il possible en un temps raisonnable ?

Pour assurer une réponse cellulaire rapide à un stimulus, le processus de recherche d'un TF pour sa petite cible d'ADN doit être extrêmement efficace et a sans doute été optimisé au cours de l'évolution des espèces.

Nous avons développé la théorie du petit trou pour étudier les propriétés d'un processus de recherche de TF [13-15]. Des expériences ont montré qu'un TF cherche son promoteur ADN en alternant des phases de diffusion libre (mouvement Brownien en trois dimension) dans le noyau et des phases de liaison à l'ADN et effectue un processus de diffusion en une dimension le long de l'ADN, mécanisme appelé diffusion facilitée.

La diffusion dans le noyau est rapide, cependant, un TF ne peut pas reconnaître sa cible s'il n'est pas lié à l'ADN. Ainsi, pour trouver sa cible, le TF doit se lier et glisser le long de l'ADN. Le problème ici est qu'une telle recherche 1D le long de l'ADN n'est pas efficace pour deux raisons: premièrement, en raison des interactions entre le squelette de l'ADN et le TF, le glissement le long de l'ADN est lent; Deuxièmement, lorsqu'il effectue une diffusion aléatoire le long de l'ADN, un TF se rendra plusieurs fois au même emplacement de l'ADN, ce qui rendra la recherche inefficace. La question est donc de savoir quelle est la meilleure stratégie d'échange entre la diffusion 3D dans le noyau et la diffusion 1D le long de l'ADN [12]?

Par analogie, considérons un problème de livraison dans Paris: un robot doit livrer un colis dans une maison spécifique, par exemple à l'Elysée. Le robot a bien la clé de la porte de la maison, mais il ne sait pas où elle se trouve. Voici le mode de recherche : le robot peut soit aller de maison de proche en proche, soit décider de faire un grand déplacement pour essayer une maison lointaine. Une stratégie pour le robot pourrait être de simplement passer de porte en porte, et de tenter d'ouvrir la porte avec sa clé jusqu'à ce qu'il trouve la maison. Cependant, cette stratégie de livraison simple n'est pas efficace, car essayer d'ouvrir toutes les portes prend du temps. De plus, le robot n'ayant aucune mémoire sur son passé et basculant de manière aléatoire entre les portes voisines, il tentera souvent d'ouvrir une même porte qu'il avait déjà tenté auparavant, ce qui rend la livraison très inefficace. Pour réduire cette redondance, le robot peut décider de basculer au bout d'un certain temps aléatoire du mode vérification vers un mode de grand déplacement où il saute de façon aléatoire (marche aléatoire) dans la rue, mais n'essaie pas d'ouvrir des portes. Parce qu'un tel déplacement dans la rue est rapide, le robot se déplace rapidement dans un quartier différent où il n'a peut-être pas encore cherché. Cependant, le robot ne devrait pas passer trop de temps dans ce mode de déplacement, car il ne peut pas trouver la maison dans ce mode. Enfin, le robot peut également prendre le métro qui l'amène dans un quartier aléatoire qui peut être éloigné de sa zone actuelle.

Ainsi, comparé à un déplacement dans la rue, prendre le métro réduit fortement la redondance de sa recherche, mais l'inconvénient est que prendre le métro prend plus de temps qu'un déplacement dans la rue. Ainsi, étant donné le nombre de maisons parisiennes, le temps nécessaire pour contrôler une maison individuelle, le temps nécessaire pour effectuer des

déplacements dans la rue et en métro : Y a-t-il une stratégie optimale entre les trois modes de recherche qui minimise le délai de livraison ?

Nous avons étudié un processus de recherche de TF avec ces transitions entre trois états en utilisant la théorie du petit trou (Fig. 4d). Nous avons identifié une stratégie optimale et calculé le temps de recherche moyen. Les modèles précédents avec seulement deux modes de recherche (sans état de déplacement le long de l'ADN) prédisaient que pour une recherche optimale, un TF devait passer le même temps à diffuser dans le noyau par rapport à un glissement le long de l'ADN. Cependant, des données expérimentales [13] et théorique [13-14] ont montré qu'un TF passe plus de temps lié à l'ADN par rapport à une diffusion libre dans le noyau. Avec notre modèle à 3 états, nous avons montré qu'une recherche optimale nécessite que le TF passe plus de temps lié à l'ADN par rapport à une diffusion libre dans le noyau. Nous avons également montré qu'une conformation d'ADN enroulée accélère la recherche, contrairement à l'intuition. Et avec notre scénario de saut d'état, un temps de recherche rapide peut être obtenu également en présence de fortes interactions entre le TF et l'ADN qui empêchent le déplacement le long de l'ADN.

En conclusion, la théorie du petit trou permet de mieux comprendre la dynamique de régulation des gènes dans le noyau, mais aussi la dynamique de la chromatine (voir aussi [11]).

5-La théorie du petit trou, l'électro diffusion où la mémoire se trouve peut-être codée.

Certaines cellules comme les neurones du cerveau ont la propriété de pouvoir générer une différence de potentiel électrique entre le monde extérieur et leur univers intérieur (Fig. 5). C'est grâce à l'énergie fournie qu'elles produisent à partir du glucose et de l'oxygène que des pompes localisées à la surface des cellules vont faire sortir des ions sodiques et faire entrer des ions potassiques.

A quoi cela sert-il ? A faire circuler et moduler un signal électrique porteur d'une information, bien qu'on ne sache pas toujours quel en est le code. Mais comment le signal électrique est-il modulé ? Où se trouve encodée la mémoire à ce niveau ? Serait-il possible que la forme à la fois stéréotypée d'une épine dendritique (Fig. 5a) et d'une synapse bien que modulée par des petites variations géométriques, joue le rôle de stockage élémentaire de la mémoire ? Comment aborder cette question ? Une façon de la faire serait d'utiliser le principe de conversion du courant électrique en voltage : C'est en effet la base du hardware électronique et des microprocesseurs que de trouver les bons assemblages de composantes pour calculer et stocker. La difficulté dans la neurobiologie est de trouver un paradigme mental pour aborder cette question. On sait depuis Benjamin Franklin que la géométrie et l'électricité sont liées, ce qui lui a permis de définir une forme pour canaliser le tonnerre (charge électrique en mouvement dans une solution). Donc des charges en mouvements devraient pouvoir interagir avec la géométrie, mais comment ?

La théorie de Hodgkin-Huxley énoncée dans les années 1950, décrit la propagation d'un grand saut de voltage d'un point vers un autre d'un axone, ressemblant à un câble électrique rectiligne, mais cette théorie ne tient pas compte du mouvement des charges électriques. Nous avons donc pensé que le mouvement des charges dans un électrolyte, comme des piles, pouvait

être influencé par la géométrie. En particulier la conversion du courant en voltage n'est plus simplement la loi d'ohm $U=RI$, mais plus complexe.

Cette question peut trouver une formulation dans les équations de l'électro-diffusion, équation de Poisson couplée aux équations de diffusion (Fokker-Planck). Alors, comme ce sont des équations aux dérivées partielles, la solution dépend crucialement de la géométrie, ce qui nous montre que nous sommes bien sur une voie qui permet d'intégrer la géométrie dans la solution de la conversion courant-Voltage. Un élément supplémentaire est que le courant entre et sort par des petits trous. Nous voilà donc ramenés aux petits trous. Un dernier élément intéressant est que nous ne pensons pas qu'il y ait électro-neutralité (c'est-à-dire, autant de charges positives que négatives à tout moment) dans des petits domaines de l'ordre de la dizaine-centaine de nanomètres. Cette propriété permet de propager beaucoup plus loin que prévu l'amplitude du champ électrique porté par la charge électrique. Pour une charge positive entrante, le neurone fait ensuite ressortir d'autres charges positives pour ne pas exploser.

En conclusion, la théorie de l'électro-diffusion en neurobiologie n'est qu'à son début pour deux points de vues : celui expérimental avec l'avènement des pipettes avec un bout de quelques nanomètres (Fig 5b-c) et les molécules luminescentes sensibles au voltage qui vont permettre dans le futur de mesurer le voltage de nano et microdomains des cellules vivantes. Mais pour interpréter ces mesures, il faudra, en même temps, un second point de vue : celui de la théorie de l'électro-diffusion pour reconstruire (Fig. 5d-e), calculer et prédire les changements de voltage due à la géométrie. Peut-être que cette théorie nous permettra de mieux comprendre le lien entre la géométrie des épines et le codage sous-cellulaire de la mémoire.

6-Le petit trou pour la sortie extrême : Pourquoi faut-il de la redondance (beaucoup de copie du même objet) : est-ce vraiment du gâchis ?

Pourquoi faut-il tant de spermatozoïdes pour fertiliser un œuf [6,8]? Pourquoi une tortue a-t-elle tant de progénitures ? Pourquoi faut-il libérer des milliers de neurotransmetteurs pour réussir la communication synaptique entre deux neurones ? Une extension de la théorie du petit trou, dite « Théorie extrême du petit trou » permet de donner une réponse à ces questions de façon surprenante : ces grands nombres sont nécessaires et ne doivent pas être vus comme du gâchis de la nature : Sans cette grande redondance (beaucoup de reproductions du même objet), la fonction biologique ne pourrait pas être possible, car elle requiert des événements rares qui doivent *être* pris en compte ? et ce, de façon incontournable. Autrement dit, la fonction biologique et l'émergence de la vie a dû mettre en place des stratégies pour contourner le hasard. Une solution de la nature est la *redondance* ou la surabondance.

7- Conclusion

Contrairement à ce que l'on pourrait penser, la biologie à petite échelle n'est pas décrite par les lois de la physique classique ou quantique *uniquement*, mais surtout par la sélection de certains comportements, et les lois sont plutôt celles de la statistique extrême : en effet, beaucoup de processus biologiques ne sont pas accomplis par le comportement moyen décrit par des trajectoires elles-mêmes décrites par la mécanique Hamiltonienne, mais par le comportement extrême du premier qui va arriver à une destination inconnue pour lui à l'avance, comme le premier spermatozoïde qui arrive à un ovule : Sa trajectoire, presque déterministe est très différente de ceux qui arriveront bien plus tard. Ceci est aussi vrai pour les trajectoires des premiers ions ou molécules qui trouvent leurs cibles cachées dans les phénomènes de transduction du signal, comme l'arrivée de premiers messagers chimiques qui ouvriront les

canaux ioniques ou les premiers ions calcium qui arriveront à un récepteur localisé sur des organelles, qui permettront de générer une avalanche et donc d'amplifier le signal. Ces exemples montrent bien que les lois d'activation de la biologie ne sont pas décrites par les comportements moyens, de lois de la physique classique, quantique ou n'importe quelle théorie qui décrit un seul objet à la fois, mais par les lois de statistiques du plus rapide parmi des copies identiques du même objet.

A noter, qu'il n'existe pas encore de principe général [7] pour trouver les trajectoires des plus rapides, qui dépendent de la physique du mouvement de chaque particule individuelle, mouvement dans lequel il y a forcément une composante aléatoire. Sans cette composante, il n'y aurait pas de place pour le choix qui fait qu'un premier puis un second vont émerger. Ce changement de point de vue [8] n'est pas simplement valable à petite échelle. Les lois des comportements extrêmes sont aussi celles qui définissent les vies des individus particuliers. Mais peut-être qu'une composante génétique est ici à considérer. Une dernière chose que cette théorie nous a appris *est* que le nombre d'individus est un élément clé pour définir le temps du premier. Par analogie avec les spermatozoïdes, pour chercher *un but ou* une cible cachée, qui sera trouvée par le premier, une trajectoire particulière est sélectionnée et dans ce cas, le nombre de participants est déterminant pour définir le temps de l'action.

La théorie du petit trou nous a montré qu'il y a des structures géométriques qui contrôlent le temps de résidences, comme les petits trous eux-mêmes ou *les longs cous...* que nous avons appelés les contrôleurs du temps. Ils définissent les temps de transmission de l'information portée par les flux ioniques des premiers ions qui arrivent à leur cible. Cette théorie géométrique de l'aléatoire est un outil de compréhension du monde sous-cellulaire à partir du niveau moléculaire.

A l'heure où les réformes scolaires privent les élèves de pouvoir suivre en même temps des cours de maths, de bio, informatique, de physique et de chimie, ces mêmes combinaisons de disciplines sont en train de permettre l'émergence d'une science computationnelle et d'une biologie mathématique qui ont déjà permis un changement de point de vue [9]: ces combinaisons ont permis de trouver des structures géométriques dans les grands jeux de données et produisent même de nouveaux théorèmes, mais surtout des découvertes en biologie. Ces nouvelles approches permettent aussi d'élaborer des simulations numériques relativement simples de systèmes très complexes comme le mouvement des ions dans les neurones. Ces méthodes ne sont pas simplement des jouets théoriques, mais elles prédisent l'organisation du vivant au niveau moléculaire-cellulaire, comme ce fut le cas récemment pour la *distribution* des récepteurs sur les organelles de la cellule [10]. La modélisation mathématique couplée à l'analyse de données en biologie est mûre pour devenir une discipline incontournable dans l'investigation du vivant. Il est peut-être encore temps de donner le bagage nécessaire pour former une nouvelle génération dès le plus jeune âge, surtout avant et pendant la puberté, pour que cette génération puisse ramener un jour la France sur le podium des premiers producteurs de savoir dans toutes les disciplines scientifiques et surtout dans celle émergeant entre les mathématiques et la biologie.

Remerciements : Nous tenons à remercier Daniel Duprez pour ces commentaires critiques, qui ont occupé les premiers jours de sa 90ème année, avec la même vivacité qu'il y a dix ans.

FIGURES:

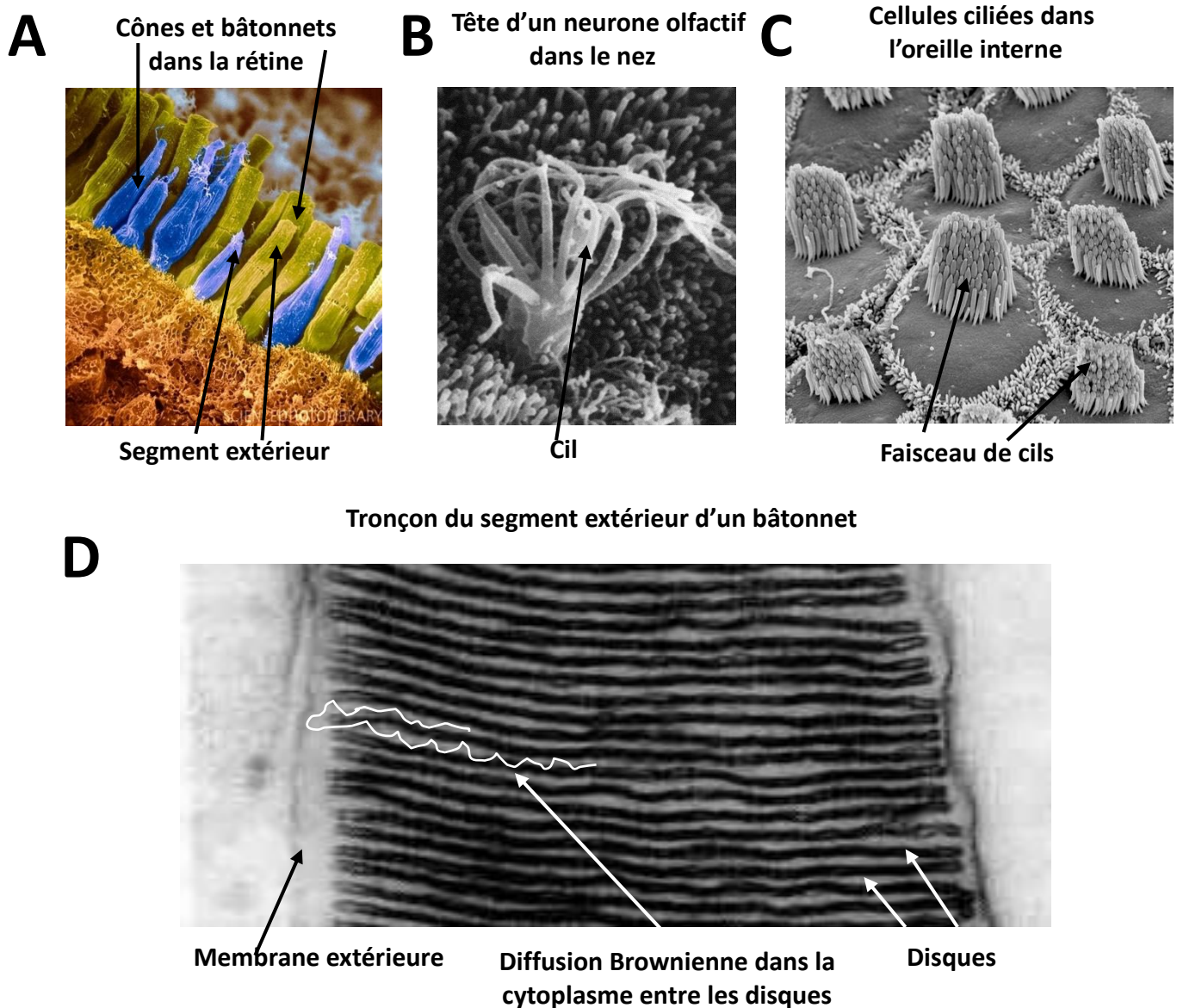


Figure 1: (A)

(Source: Science Photo Library); Le segment extérieur et le lieu où les photons sont absorbés. **(B)**

(de Morrison & Costanzo (1990)); **(C)**

erne (image de asbmb.org). **(D)** Exemple d'une trajectoire Brownienne pour un ion calcium ou une molécule cGMP évoluant entre deux disques du segment externe d'un bâtonnet

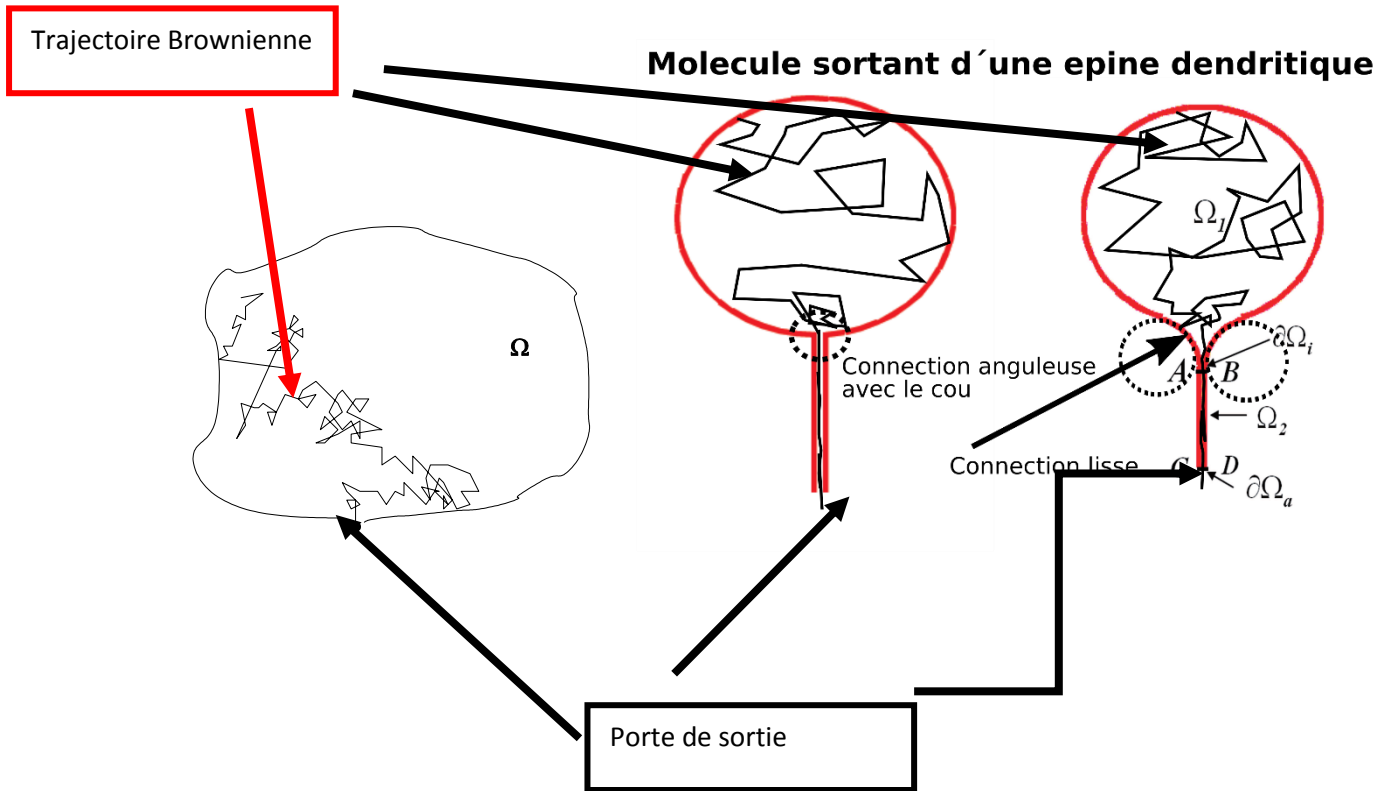


Figure 2: Trois exemples de sortie par un petit trou. A gauche : une trajectoire sort par une petite porte de sortie dans le domaine qui s'appelle Ω . Au milieu et à droite : sortie par un domaine avec un cou. Le temps devient plus long du aux aller-retours entre la tête et le cou a l'intersection notée AB avec de sortir en CD ou le domaine absorbant est note $\partial\Omega_a$.

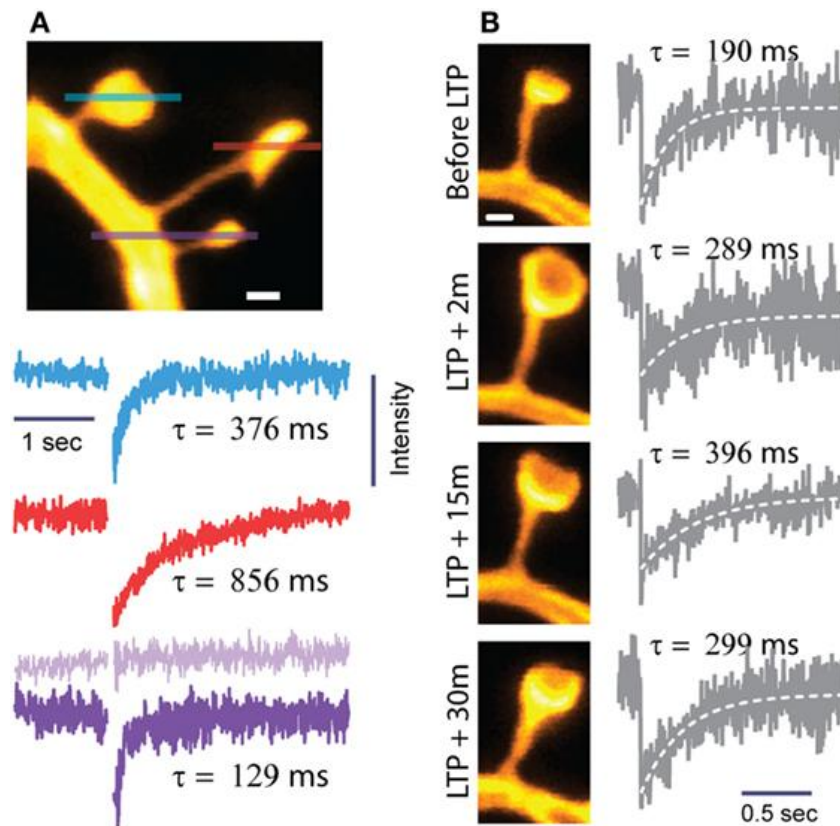


Fig3 : Exemples de temps de sortie du flot de particules aléatoires par une épine dendritique A. B montre 4 exemples après un protocole qui simule l'apprentissage au niveau d'une synapse neuronale. Le temps tau extrait à partir d'un fit d'une fonction exponentielle. Ce taux est exactement l'inverse du premier temps de passage comme explique dans la fig 2.

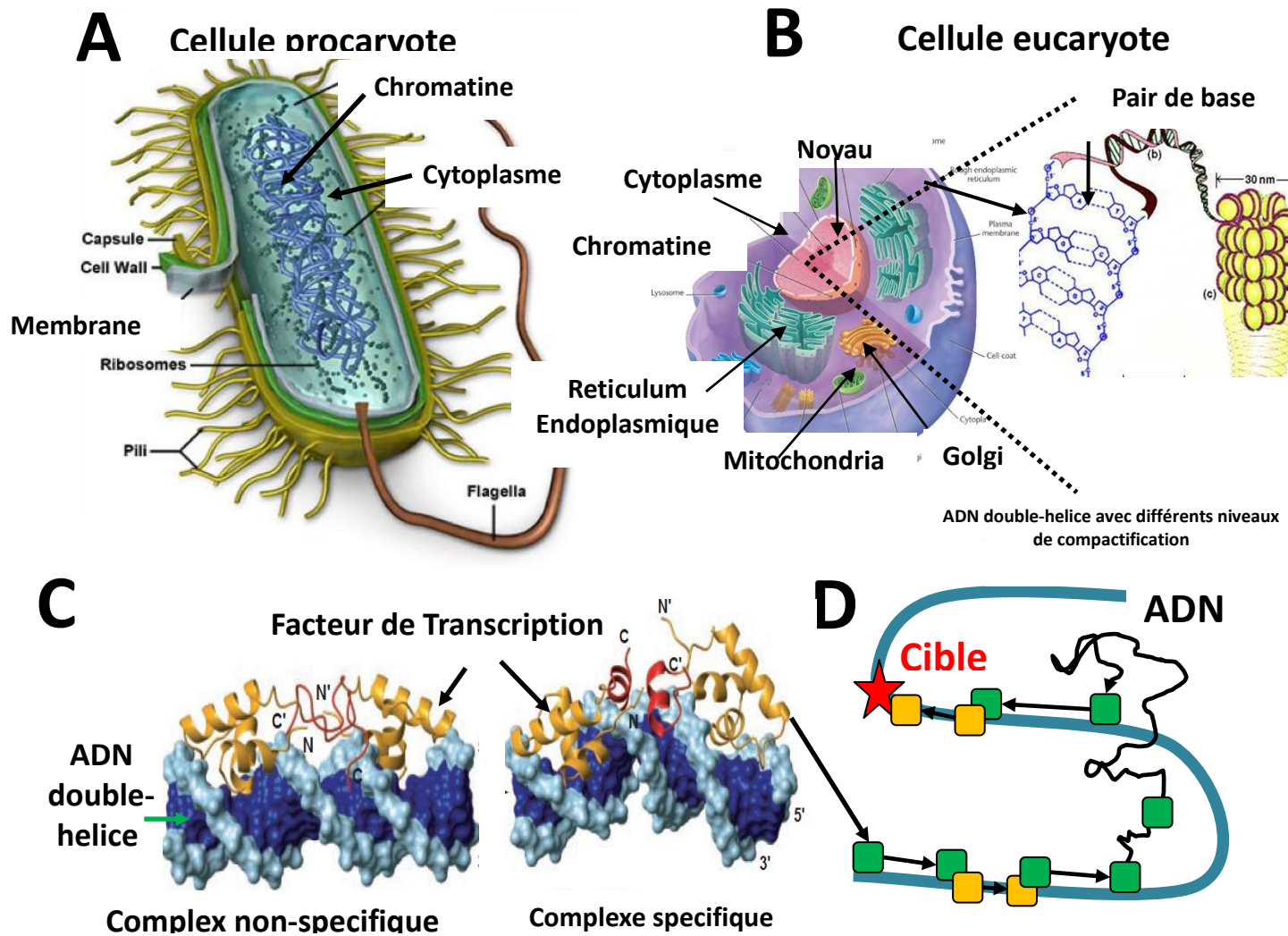


Figure 4: (A) Cellule bactérienne avec ADN librement accessible dans le cytoplasme (© 1995-2019 par Michael W. Davidson et The Florida State University). (B) Cellule eucaryote où l'ADN est confiné au noyau. Le grossissement montre la structure de l'ADN avec les différents niveaux de compactification de l'ADN (provenant de biologydictionary.net) (C) le facteur de transcription (TF) lacépresseur lié à l'ADN dans sa conformation protéique spécifique et non spécifique (d'après Kalodimos et al., Science). 2004). (D) Schéma du processus de recherche d'un TF dans une bactérie: lorsqu'il est fixé à l'ADN par sa conformation non spécifique (vert), il génère un complexe non spécifique où la diffusion le long de l'ADN est rapide. Après être passé au complexe spécifique (orange), un TF scanne et interagit fortement avec l'ADN et la diffusion est lente. Dans la conformation non spécifique, un TF peut également se détacher de l'ADN et diffuser librement à l'intérieur du noyau. Après un certain temps aléatoire, il se lie à l'ADN et poursuit le processus de recherche le long de l'ADN. Ce cycle se répète jusqu'à ce qu'il atteigne et se lie à la cible ADN.

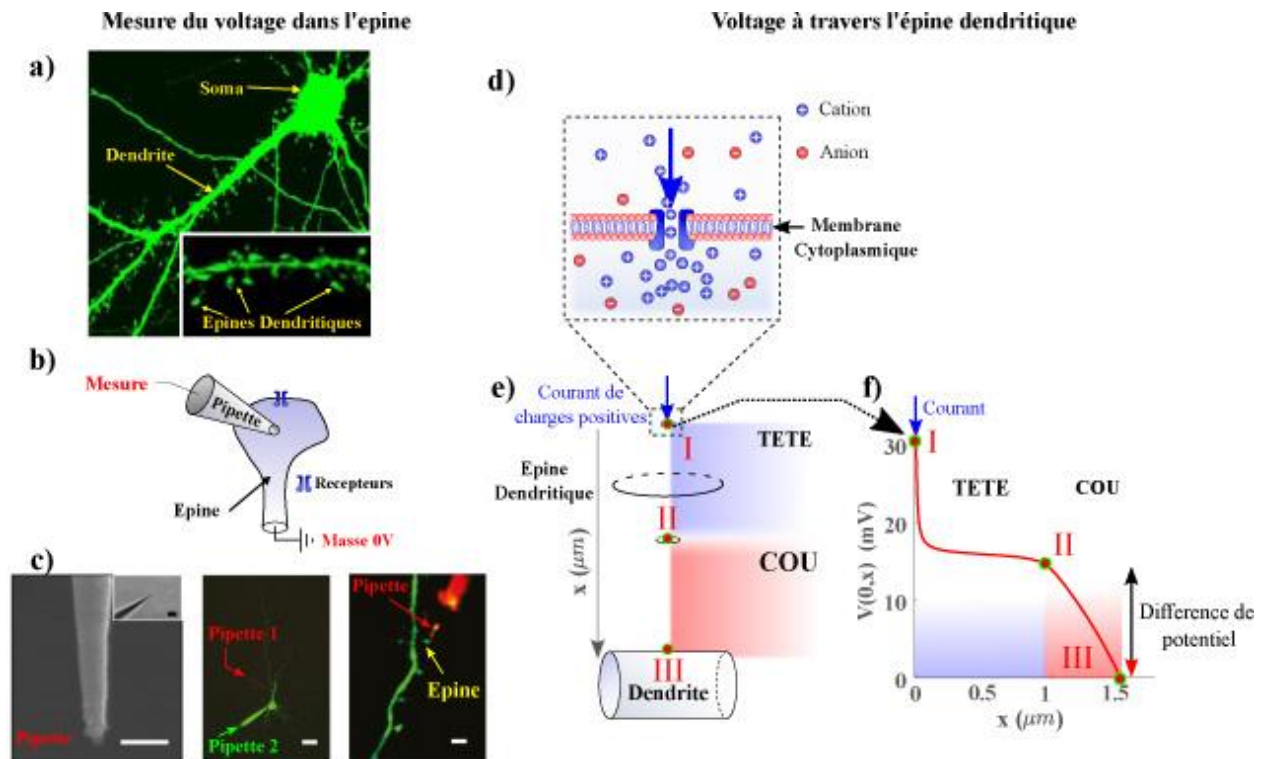


Figure 5 : Mesure et calcul du Voltage dans une épine dendritique : **a)** Image d'un neurone obtenue par microscopie à fluorescence, avec dans l'encart un tronçon du dendrite recouvert d'épines. **b)** Schéma explicatif de la mesure du voltage dans la tête d'une épine dendritique via une nano-pipette. **c)** Trois panels successifs montrant : le bout d'une nano-pipette (échelle : 100nm), l'enregistrement simultané de l'épine (pipette 1) et du soma (pipette 2), échelle = 20 μm , et un zoom sur la région où est mesuré le voltage dans la tête de l'épine (échelle : 1.5 μm). **d)** Schéma explicatif de l'influx de cations à travers un canal ionique situé dans la membrane cellulaire de la tête de l'épine. **e)** Modélisation de l'épine dendritique en deux compartiments : la tête sphérique et le cou cylindrique. Les disques rouges désignent : I le sommet de la tête, II la jonction tête-cou, et III la jonction cou\dendrite. **f)** Distribution du voltage le long de l'axe de l'épine dendritique suite à l'application d'un courant cationique au sommet de la tête obtenue par résolution numérique des équations de Poisson-Nernst-Planck. Les régions I, II, et III sont les mêmes que celles décrites dans le panel e).

References

- 1-D Holcman, Z Schuss [Stochastic narrow escape in molecular and cellular biology](#) Analysis and Applications. Springer, New York 2015
- 2-D Holcman, Z Schuss [Asymptotics of Elliptic and Parabolic PDEs](#) Springer International Publishing 2018
- 4- Holcman, D., and Z. Schuss. "Escape through a small opening: receptor trafficking in a synaptic membrane." *Journal of Statistical Physics* 117.5-6 (2004): 975-1014.
- Singer, A., Schuss, Z., & Holcman, D. (2006). Narrow escape, Part II: The circular disk. *Journal of statistical physics*, 122(3), 465-489.
- Singer, A., Schuss, Z., & Holcman, D. (2006). Narrow escape, Part III: Non-smooth domains and Riemann surfaces. *Journal of statistical physics*, 122(3), 491-509.
- Singer, A., Schuss, Z., Holcman, D., & Eisenberg, R. S. (2006). Narrow escape, part I. *Journal of Statistical Physics*, 122(3), 437-463.
- Holcman, D., & Schuss, Z. (2012). Brownian motion in dire straits. *Multiscale Modeling & Simulation*, 10(4), 1204-1231.
- Singer, A., Schuss, Z., & Holcman, D. (2008). Narrow escape and leakage of Brownian particles. *Physical Review E*, 78(5), 051111.
- Pillay, S., Ward, M. J., Peirce, A., & Kolokolnikov, T. (2010). An asymptotic analysis of the mean first passage time for narrow escape problems: Part I: Two-dimensional domains. *Multiscale Modeling & Simulation*, 8(3), 803-835.
- Cheviakov, A. F., Ward, M. J., & Straube, R. (2010). An asymptotic analysis of the mean first passage time for narrow escape problems: Part II: The sphere. *Multiscale Modeling & Simulation*, 8(3), 836-870.
- 4-D Holcman, Z Schuss [100 years after Smoluchowski: stochastic processes in cell biology](#) Journal of Physics A: Mathematical and Theoretical 50 (9), 1-45
- 5- D Holcman, [Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology](#) Springer 2017
- 6-K Reynaud, Z Schuss, N Rouach, D Holcman [Why so many sperm cells?](#) Communicative & integrative biology 8 (3), e1017156 2015
- 7-J Reingruber, D Holcman, GL Fain [How rods respond to single photons: Key adaptations of a G-protein cascade that enable vision](#) Issue: BioEssays, 37 (11) 2015
- 8-Z Schuss, K Basnayake, D Holcman [Redundancy principle and the role of extreme statistics in molecular and cellular biology](#) Physics of life reviews 2019
- K Basnayake, D Holcman, Fastest among equals: A novel paradigm in biology: Reply to comments: Redundancy principle and the role of extreme statistics in molecular and cellular biology, Physics of Life Reviews 2019

Basnayake K, Mazaud D, Bemelmans A, Rouach N, Korkotian E, Holcman D., Fast calcium transients in dendritic spines driven by extreme statistics. PLoS Biol. 2019;17(6):e2006202. doi: 10.1371/journal.pbio.2006202.

9-D Holcman, Z Schuss, [Commentary New mathematical physics needed for life sciences](#)
Physics Today 69 (1), 10-12 2016

10-Schuss, Z. (2009). *Theory and applications of stochastic processes: an analytical approach* (Vol. 170). Springer Science & Business Media.

Et

Schuss, Z. ,Theory and Applications of Stochastic Differential Equations (Wiley Series in Probability and Statistics - Applied Probability and Statistics Section) 1980.

11- Amitai, A., & Holcman, D. (2017). Polymer physics of nuclear organization and function. *Physics Reports*, 678, 1-83.

Et

Amitai, A., Kupka, I., & Holcman, D. (2012). Computation of the mean first-encounter time between the ends of a polymer chain. *Physical review letters*, 109(10), 108302.

12-Reingruber, J., & Holcman, D. (2009). Gated narrow escape time for molecular signaling. *Physical review letters*, 103(14), 148102

13- Malherbe, G., and D. Holcman. "The search kinetics of a target inside the cell nucleus." *arXiv preprint arXiv:0712.3467* (2007).

Elf, J., Li, G. W., & Xie, X. S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, 316(5828), 1191-1194.

Malherbe, G., and D. Holcman. "The search for a DNA target in the nucleus." *Physics Letters A* 374.3 (2010): 466-471.

14-Reingruber, J., & Holcman, D. (2011). Transcription factor search for a DNA promoter in a three-state model. *Physical Review E*, 84(2), 020901.

15-Cartailler, J., & Reingruber, J. (2015). Facilitated diffusion framework for transcription factor search with conformational changes. *Physical biology*, 12(4), 046012.

Boite 1 pour les curieux : formulation mathématique de la théorie du petit trou :

Pour un mouvement aléatoire soumis à un champ de force F et un mouvement Brownien w , la vitesse suit la loi de Smoluchowski :

$$\dot{x} + \frac{1}{\gamma} F(x) = \sqrt{2D} \dot{w}, (*)$$

ou $F = -\nabla \phi(x)$, γ est le coefficient de friction. Alors il est possible de trouver l'équation pour le temps moyen τ ou la durée de séjour moyenne d'une particule aléatoire suivant la loi (*) dans un domaine délimité, avant que celle-ci ne s'échappe par une petite fenêtre absorbante située à sa frontière limite. La partie restante du bord est supposée réfléchir le mouvement de la particule. La réflexion peut représenter une barrière à fort potentiel sur la limite ou un obstacle physique. L'ouverture peut représenter un couloir étroit dans la barrière ou un défaut de l'obstacle physique (voir la figure 1).

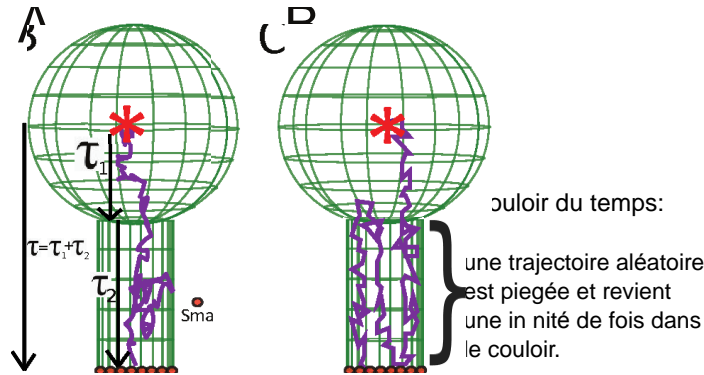


Figure 1 : Trajectoire aléatoire sortant par un petit trou, localisée à la fin d'un long cou.

Le temps de sortie peut être estimé de manière asymptotique dans la limite $\varepsilon = \frac{|\partial\Omega_a|}{|\partial\Omega|} \ll 1$. La

densité de probabilité des trajectoires de la figure 1 (ci-dessus) est la probabilité par unité de volume (surface) de trouver la particule aléatoire au point situé à un moment précédant sa sortie.

$$\frac{\partial p_\varepsilon(x,t)}{\partial t} = D\Delta p_\varepsilon(x,t) + \nabla \cdot p_\varepsilon(x,t) \nabla \phi(x),$$

Le PDF est solution de l'équation de Fokker-Planck

avec la condition initiale $p_\varepsilon(x,0) = \rho_0(x)$, (par exemple, pour une distribution uniforme

$\rho_0(x) = 1/|\Omega|$ ou, lorsque la molécule est initialement située en position y $\rho_0(x) = \delta(x-y)$, ou

δ est la fonction de Dirac. Les conditions aux limites mixtes Dirichlet-Neumann sont

$$p_\varepsilon(x,t) = 0, \quad \text{for } x \in \partial\Omega_a, t > 0$$

$$D \frac{\partial p_\varepsilon(x,t)}{\partial n} + \frac{p_\varepsilon(x,t)}{\gamma} \frac{\partial \phi(x)}{\partial n} = 0, \quad \text{for } x \in \partial\Omega_r, t > 0,$$

Alors la fonction $u_\varepsilon(y) = \int_0^\infty dx \int_\Omega p_\varepsilon(x,t|y) dt$, est le temps moyen et satisfait l'équation aux dérivées partielles

$$D\Delta u_\varepsilon(y) - \nabla \phi(y) \cdot \nabla u_\varepsilon(y) = -1, \quad \text{for } y \in \Omega$$

$$u_\varepsilon(y) = 0, \quad \text{for } y \in \partial\Omega_a \quad (**)$$

$$\frac{\partial u_\varepsilon(y)}{\partial n} = 0, \quad \text{for } y \in \partial\Omega_r.$$

Enfin, la probabilité de trouver la particule dans le domaine au temps t est

$$S_\varepsilon(t) = \int_{\Omega} p_\varepsilon(x, t) dx, \approx e^{-\lambda_0(\varepsilon)t} \text{ pour } t \gg \frac{1}{\lambda_1(\varepsilon)}, \quad \lambda_i(\varepsilon) \text{ sont les valeurs propres de l'opérateur}$$

(**)

Les méthodes asymptotiques permettent de trouver la solution des équations (**), qui dépend de la dimension (2 ou 3) et de la géométrie locale à proximité de la petite ouverture [1].

Boîte 2 : Des formules explicites de temps de sortie par un petit trou

Lorsque $\varepsilon = \frac{|\partial\Omega_a|}{|\partial\Omega|} \ll 1$, le temps de sortie pour un mouvement Brownien est donné par les formule suivante

$$u_\varepsilon(y) = \begin{cases} \frac{|\Omega|}{\pi D} \ln \frac{1}{\varepsilon} + O(1) & \text{for } n = 2 \\ \frac{|\Omega|}{4aD} 1 + o(1) & \text{for } n = 3, \end{cases}$$

Où D est le coefficient de diffusion, $|\Omega|$ est le volume (dim3) ou la surface (dim2) du domaine, a représenté le petit rayon d'un disque situé à la surface du domaine. A l'ordre dominant, le temps de sortie $u_\varepsilon(y)$ ne dépend pas de la position initiale y , à l'exception d'une petite couche limite proche, en raison de la forme asymptotique [1].

En dimension 2, le terme de premier ordre est important, car, par exemple, si, le second terme d'est comparable au terme principal. Le second terme peut être trouvé pour un disque circulaire de rayon a et quand la particule commence au centre, le temps est

$$E[\tau | x(0) = 0] = \frac{|\Omega|}{\pi D} \left[\log \frac{1}{\varepsilon} + \log 2 + \frac{1}{4} + O(\varepsilon) \right].$$

Si la fenêtre absorbante est située à un angle α , alors le temps devient

$$\bar{\tau} = \frac{|\Omega|}{\alpha D} \left[\log \frac{1}{\varepsilon} + O(1) \right].$$

Plus surprenant, près d'un coup dans un domaine à deux dimensions, le temps d'échappement augmente algébriquement plutôt que logarithmique: dans le domaine délimité par une fenêtre

situé entre deux cercles, alors $\bar{\tau} = \frac{\pi |\Omega|}{D\sqrt{\varepsilon/R}} \left(1 + O\left(\sqrt{\frac{\varepsilon}{R}}\right) \right)$, où R est le rayon de courbure.

Lorsqu'il y a deux petits trous circulaires de rayon $2\varepsilon, 2\delta$ séparés par la distance Δ entre leurs centres, alors

$$\bar{\tau} = \frac{|\Omega|}{\pi D \left(\log \frac{1}{\varepsilon} + \log \frac{1}{\delta} \right)} \times \frac{\left(\log \frac{1}{\delta} \log \frac{1}{\varepsilon} - \log \Delta \right)^2}{1 + 2 \frac{\log \Delta}{\log \frac{1}{\delta} + \log \frac{1}{\varepsilon}}}$$

En dimension 3, la formule peut se compléter et un terme nouveau apparaît qui représente la courbure moyenne au centre

$$\bar{\tau} = \frac{|\Omega|}{4aD \left[1 + \frac{L(0) + N(0)}{\pi} a \log a + o(a \log a) \right]},$$

La forme géométrique du petit trou influence le temps de sortie : dans le cas d'un trou en forme d'un ellipse d'axe a et b, le temps devient

$$\bar{\tau} \approx \frac{|\Omega|}{2\pi Da} K(e) + \text{termes inconnus}$$

ou $K(e) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - \frac{a^2 - b^2}{b^2} \sin^2 \theta}}$ est l'intégrale elliptique. Dans le cas d'ellipsoïde de révolution,

des formules similaires ont été trouvées $\tau = \frac{|V|}{4a_1 D} K(e, f)$ avec

$$K(e, f) = \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{(1 - e^2 \sin^2 \theta)(1 + f^2 \tan^2 \theta)}} \text{ et } f = a_3/a_1, a_2 \geq a_1, e = \sqrt{1 - (a_1/a_2)^2}$$

[1-1b]. Pour une fenêtre en forme d'épine avec $a_2 = a_1$

$$\tau = \frac{|V|}{2\pi a_1 D} \begin{cases} \frac{\arccos f}{\sqrt{1 - f^2}}, & f \leq 1 \\ \frac{\arg \cosh f}{\sqrt{f^2 - 1}}, & f \geq 1 \end{cases}$$

Pour f grand, $\tau \approx \frac{|\Omega|}{2\pi a_1 D} \frac{\ln(2f)}{f}$.

Lorsque de long couloirs, comme le cou d'une épine dendritique (Fig. 3) sépare le milieu où évolue la particule aléatoire et la porte de sortie, alors le temps de recherche devient beaucoup plus long que prévu à cause des nombreux aller-retour entre le couloir et le domaine. Ces couloirs sont appelés couloir du temps car la particule aléatoire peut retourner en moyenne une infinité de fois dans le domaine avant de trouver la cible. C'est comme si elle recommençait à la même position une infinité de fois. Dans ce cas les formules du petit trou deviennent :

$$\tau_{3\text{dim}} = \begin{cases} \frac{|\Omega_1| L}{\pi a^2 D} + \frac{|\Omega_1|}{4aD} \left[1 + \frac{a}{\pi R} \log \frac{R}{a} \right] + \frac{L^2}{2D} & \text{pour une boule avec un cou} \\ \frac{|\Omega_1| L}{\pi a^2 D} + \frac{1}{\sqrt{2}} \left(\frac{R_c}{a} \right)^{3/2} \frac{|\Omega_1|}{R_c D} + \frac{L^2}{2D} & \text{pour un domaine quelconque (liaison avec un cusp)} \end{cases}$$

Où L est la longueur du cou, R_c le rayon de courbure, D le coefficient de diffusion, a le rayon du couloir, R le rayon de la boule où évolue la particule.

Une question ouverte : Quelle formule pour une fenêtre de la forme d'un rectangle ou un polygone régulier (100 Euros de récompense pour celle/celui qui résoudra la question). Dans le cas où il y a des forces effectives, les formules se modifient. Elles ont été dérivées pour modéliser les premières étapes de l'infection viral dans les cellules [1-4] ou le mouvement des vésicules.

References de la boîte 2:

[1] Holcman, D. and Schuss, Z., 2015. Stochastic narrow escape in molecular and cellular biology. *Analysis and Applications*. Springer, New York.

[1b] Reingruber, Jürgen, Enrique Abad, and David Holcman. "Narrow escape time to a structured target located on the boundary of a microdomains." *The Journal of chemical physics* 130, no. 9 (2009): 094909.

[2] Lagache, T, and D. Holcman. "Quantifying intermittent transport in cell cytoplasm." *Physical Review E* 77, no. 3 (2008): 030901.

[3] Holcman, D., 2007. Modeling DNA and virus trafficking in the cell cytoplasm. *Journal of Statistical Physics*, 127(3), pp.471-494.

[4] Lagache, T., and D. Holcman. "Extended narrow escape with many windows for analyzing viral entry into the cell nucleus." *Journal of Statistical Physics* 166, no. 2 (2017): 244-266.