Facets on the convex hull of $d$-dimensional Brownian and Lévy motion

Julien Randon-Furling
SAMM (EA 4543), Université Paris-1 Panthéon-Sorbonne, Centre Pierre Mendès-France, 90 rue de Tolbiac, 75013 Paris, France

Florian Wespi
IMSV, Universität Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
(Dated: November 16, 2018)

For stationary, homogeneous Markov processes (viz., Lévy processes, including Brownian motion) in dimension $d \geq 3$, we establish an exact formula for the average number of $(d-1)$-dimensional facets that can be defined by $d$ points on the process’s path. This formula defines a universality class in that it is independent of the increments’ distribution, and it admits a closed form when $d = 3$, a case which is of particular interest for applications in biophysics, chemistry and polymer science.

We also show that the asymptotical average number of facets behaves as $\langle F_{d}^{(d)} \rangle \sim 2 \ln (T/\Delta t)^{d-1}$, where $T$ is the total duration of the motion and $\Delta t$ is the minimum time lapse separating points that define a facet.

PACS numbers: 05.40.Fb, 05.40.Jc, 02.50.-r, 36.20.Ey
INTRODUCTION

Stochastic processes’ paths are common in the physical and natural sciences, in particular as models of actual motion [1–4], interfaces [5] or long proteins and other polymers [6–9]. Spatial characteristics of stochastic processes thus come to reflect physically meaningful variables: home ranges of animals [10], heights of surfaces [11], shapes of molecules [12, 13].

One way to describe the spatial extent of a stochastic process’s path is to examine its convex hull. That is, the smallest convex set enclosing all points visited by the process. In two dimensions, exact expressions for the values of geometric quantities such as the mean perimeter or the mean area of the convex hull have been found for Brownian motion [14–17] as well as for other types of stochastic motions [18–20], for multi-walker systems [21] and for confined configurations [22, 23]. In some cases, equivalent results exist in higher dimensional space [19, 24–28].

Results [12, 29, 30] pertaining to the shape of such random convex hulls appear scarcer though, despite their importance in biophysics: a protein’s or an antibody’s ability to play its part depends crucially on its shape [31, 32]. Convex hulls methods are central in the description of long-molecule configurations [33–37], leading to questions on the structure of these hulls, such as on their number of edges (in 2D) or facets (in 3D) and on the distribution of distances between the molecule and its convex hull’s edges or facets [38, 39].

Only recently, the average number of edges on the boundary of the convex hull of a vast class of planar random walks was computed. Remarkably, this was found to have a universal form: it grows logarithmically with time [40], whatever the distribution of the infinitesimal increments along the process’s path – as long as these increments are independent and identically distributed (i.e. as long as the process is what is called a Lévy process).

In this paper, we show that the universality observed for the average number of edges in two dimensions extends to the average number of facets in $d \geq 3$ dimensions (Fig.1). We establish an exact, integral formula for the average number of facets on the convex hull of a vast class of Lévy processes in dimension $d$. When $d = 3$, this formula admits a closed form,

$$\langle F_T^{(3)} \rangle = 2 \left[ \ln \left( \frac{T}{\Delta t} \right) \right]^2 + 4 \left\{ \ln \left( \frac{T}{\Delta t} \right) \ln \left( 1 - \frac{\Delta t}{T} \right) - \text{Li}_2 \left( 1 - \frac{\Delta t}{T} \right) + \pi^2 / 12 \right\},$$

where $\text{Li}_2$ is the dilogarithm function, $T$ is the total duration of the process and $\Delta t$ is a cut-off parameter representing an arbitrary minimal time separation between two extreme points on the process’s path (see details in section 2). Asymptotics are obtained from the general $d$-dimensional formula,

$$\langle F_T^{(d)} \rangle \sim 2 \left[ \ln \left( \frac{T}{\Delta t} \right) \right]^{d-1},$$

which generalizes the two-dimensional case, where the average number of edges on the boundary of the convex hull was shown in [40] to be given by $2 \ln \left( \frac{T}{\Delta t} \right)$.

For the sake of clarity, we first examine the 3-dimensional discrete-time case. In section 2 we then establish an integral formula in the general $d$-dimensional, continuous-time case, before deriving asymptotics and obtaining a closed form when $d = 3$ in section 3. We illustrate our results with numerical simulations.
DISCRETE-TIME CASE

In three-dimensional space, the convex hull of the successive positions of a random walker is almost surely made up of triangular facets. Following [40], we shall compute the average number of such facets by computing the probability for a triangle defined by three points on the walker’s path to form a facet of the convex hull (see Fig. 1).

Let \( S_n \) denote the position of the walker after \( n \) steps \((S_0 = 0, n \geq 3)\). Given three integers \( n_1 < n_2 < n_3 \), the triangle joining \( S_{n_1}, S_{n_2}, \) and \( S_{n_3} \) will be a facet of the walk’s convex hull if the walker always stays on the same side of the plane defined by \( S_{n_1}, S_{n_2}, S_{n_3} \). Let us choose a reference frame where the \( xy \)-plane coincides with the one defined by \( S_{n_1}, S_{n_2}, S_{n_3} \). Then the three-dimensional random walk always stays on the same side of the \( xy \)-plane if its \( z \)-coordinate is a one-dimensional random walk that: (i) starts at some value – say a positive one – and hits 0 for the first time at time \( n_1 \); (ii) performs a positive excursion away from 0 between \( n_1 \) and \( n_2 \) (i.e. it hits 0 at these times but remains positive in between); (iii) performs another positive excursion away from 0 between \( n_2 \) and \( n_3 \); and finally (iv) leaves from 0 at time \( n_3 \) and stays positive thereafter.

Relabeling \( i = n_1 \), setting \( k_1 = n_2 - n_1 \) and \( k_2 = n_3 - n_2 \), one may formally write the following formula for the mean number \( \langle F_N^{(3)} \rangle \) of facets

\[
\langle F_N^{(3)} \rangle = 2 \sum_{k_1=1}^{N-1-k_1} \sum_{k_2=1}^{N-(k_1+k_2)} m_i^{(0)} e_{i+k_1} e_{i+k_1+k_2} m_N^{(i+k_1+k_2)},
\]

with \( m_i^{(0)}, e_{i+k_1}, e_{i+k_1+k_2}, m_N^{(i+k_1+k_2)} \) standing for the probabilities of each of the four parts (i)-(iv) described above. The prefactor 2 accounts for the fact that there are two sides to every facet.

Since we are considering only random walks with independent and identically distributed jumps, we have \( e_{i+k_1} = e_{k_1}, e_{i+k_1+k_2} = e_{k_2} \) for all \( i, k_1, \) and \( k_2 \). Therefore, Eq. (2) becomes

\[
\langle F_N^{(3)} \rangle = 2 \sum_{k_1=1}^{N-1-k_1} \sum_{k_2=1}^{N-(k_1+k_2)} m_i^{(0)} m_N^{(i+k_1+k_2)}.
\]

Markovian independence allows one to stitch the parts of the \( z \)-random walk described above as pleases (because the cuts occur at special times, namely stopping times). In particular, stitching together the first and fourth parts produces a sub-walk of duration \( N - (k_1 + k_2) \). This walk starts at some positive value and hits its minimum (0 by construction) at time \( i = n_1 \). The probability associated with this sub-walk is just the probability for a random walk with \( N - (k_1 + k_2) \) steps to hit its minimum at step \( i \) (see [40] for more details). Thus interpreting the product \( m_i^{(0)} m_N^{(i+k_1+k_2)} \) in Eq. (3), one finds that

\[
\sum_{i=0}^{N-(k_1+k_2)} m_i^{(0)} m_N^{(i+k_1+k_2)} = \sum_{i=0}^{N-(k_1+k_2)} \text{Prob}(\text{Min. is hit at } i) = 1.
\]

Substituting into Eq. (3) yields

\[
\langle F_N^{(3)} \rangle = 2 \sum_{k_1=1}^{N-1-k_1} \sum_{k_2=1}^{N-k_2} e_{k_1} e_{k_2}.
\]

There remains to compute the excursion probabilities \( e_{k_1}, e_{k_2} \). As in [40], let us notice that the excursion probability for a random walk is the probability that the walk, pinned at 0 at times 0 and \( N \), visits only the positive half space. This is simply given by the probability that a bridge (that is, a random walk pinned at 0 at times 0 and \( N \)) attains its minimum at the initial step. The well-known uniform law for bridges (linked to the arcsine law for free random walks [41,43]) therefore gives simply \( e_k = 1/k \). (Note that using generalizations of the arcsine and uniform laws [44], one could treat the three-dimensional case here without resorting to a projection on one of the coordinates. We do it only for the sake of clarity.)

Finally, we obtain

\[
\langle F_N^{(3)} \rangle = 2 \sum_{k_1=1}^{N-1} \sum_{k_2=1}^{N-k_2} \frac{1}{k_1 k_2} = 2 \sum_{k_1+k_2 \leq N} \frac{1}{k_1 k_2}.
\]
In higher dimension, the same reasoning is readily carried out. Counting facets that can be defined by \( d \) points on the path of a random walk with \( N \geq d \) steps, one obtains a \((d - 1)\)-fold sum

\[
\langle F_N^{(d)} \rangle = 2^d \sum_{k_1 + \cdots + k_{d-1} \leq N \atop k_1, \ldots, k_{d-1} \geq 1} \frac{1}{k_1 \cdot k_2 \cdots k_{d-1}}.
\]  

(6)

While writing this paper, the result published recently in [45] came to our attention: this is similar to our Eq. (6). Several other results pertaining to the discrete-time case are to be found in [45], but none on the continuous-time case, which we now examine.

**CONTINUOUS-TIME CASE**

Continuous-time analogs to stationary, homogeneous random walks are known as Lévy processes [46]. They have independent, identically distributed (iid) increments and examples are Brownian motion, Cauchy processes, Lévy-Smirnov processes and stable processes [17].

In two dimensions, the mean number of edges on the convex hull of such processes’ paths exhibits a universal form, \( \langle F_T^{(2)} \rangle = 2 \ln (T/\Delta t) \), where \( T \) is the total duration of the process. The cutoff parameter \( \Delta t \) may be interpreted as the minimal time separation, between two points on the process’s path, for them to be examined as potential endpoints of an edge on the convex hull (see [40, 48] for a detailed explanation).

Given the universality observed in both the discrete-time case and the two-dimensional continuous time case, one expects a simple continuous version of Eq.(6) to hold in the \( d \)-dimensional continuous-time case.

This may be proven using the same strategy as in the previous section. In dimension \( d \), a facet will be a \((d - 1)\)-dimensional object lying in a hyperplane and defined by \( d \) points on the process’s path. One thus splits the path into \( d + 1 \) independent parts, and then computes the probability density (pdf) associated with each part before multiplying them together and integrating. When \( d = 3 \), the 4 parts and their associated pdfs are:

- from time 0 up to time \( \tau \), the path lies on one side only (say side +) of a (hyper)plane \( \mathcal{H} \), and it hits \( \mathcal{H} \) at time \( \tau \); write \( p_T^{(0)}(\tau) \) for the corresponding pdf;
- from time \( \tau \) up to time \( \tau + k_1 \), the path is an excursion away from \( \mathcal{H} \) (in side +) and is pinned on \( \mathcal{H} \) at times \( \tau \) and \( \tau + k_1 \); write \( f(\tau, \tau + k_1) \) for the pdf;
- from time \( \tau + k_1 \) up to time \( \tau + k_1 + k_2 \), the path is an excursion away from \( \mathcal{H} \) (in side +) and is pinned on \( \mathcal{H} \) at times \( \tau + k_1 \) and \( \tau + k_1 + k_2 \); write \( f(\tau + k_1, \tau + k_1 + k_2) \) for the pdf;
- from time \( \tau + k_1 + k_2 \) up to time \( T \), the path lies in side + of \( \mathcal{H} \) given that it is pinned on \( \mathcal{H} \) at time \( \tau + k_1 + k_2 \); write \( p_T^{(\tau + k_1 + k_2)}(\tau) \) for the pdf.

Formally, one obtains the following integral

\[
\langle F_T^{(3)} \rangle = 2 \int_{\Delta t}^{T-\Delta t} \int_{\Delta t}^{T-k_1} \int_0^{T-(k_1+k_2)} dk_1 \, dk_2 \, d\tau \, p_T^{(0)}(\tau) f(\tau, \tau + k_1) f(\tau + k_1, \tau + k_1 + k_2) p_T^{(\tau + k_1 + k_2)}. \tag{7}
\]

Since Lévy processes have stationary and independent increments, we have, just as in the discrete-time case, for all \( \tau, k_1 \) and \( k_2 \): \( f(\tau, \tau + k_1) = f(0, k_1) \equiv f(k_1) \) and \( f(\tau + k_1, \tau + k_1 + k_2) = f(0, k_2) \equiv f(k_2) \). Also, the product \( p_T^{(0)}(\tau + k_1 + k_2) p_T^{(\tau + k_1 + k_2)} \) can be interpreted as giving the pdf of the time \( \tau \) at which a similar process of duration \( T - (k_1 + k_2) \) attains its minimum. Thus,

\[
\int_{\tau=0}^{T-(k_1+k_2)} p_T^{(0)}(\tau) p_T^{(\tau + k_1 + k_2)} d\tau = 1, \tag{8}
\]

and the combined integral in Eq. (7) reduces to

\[
\langle F_T^{(3)} \rangle = 2 \int_{\Delta t}^{T-\Delta t} \int_{\Delta t}^{T-k_1} dk_1 \, dk_2 \, f(k_1) f(k_2). \tag{9}
\]

Lastly, \( f(k) \) may be viewed as the pdf that the sojourn time of a process on a given side of a (hyper)plane to which it is pinned (at both endpoints) is equal to the full duration \( k \). This means, \( f(k) \) is the pdf of the sojourn
time of a Lévy bridge, which is known to be uniform in higher dimensions, just as it is in dimension 1 \cite{41}. Therefore \( f(k) = 1/k \), and we obtain

\[
\langle F_T^{(3)} \rangle = 2 \int_{k_1 + k_2 \leq T} \frac{dk_1 dk_2}{k_1 k_2}.
\]

(10)

In higher dimension, Eq. (7) becomes a \( d \)-fold integral, accounting for facets defined by \( d \) points and characterised by \( d - 1 \) excursions:

\[
\langle F_T^{(d)} \rangle = 2 \int_{T - (d - 2)\Delta t} T \int_{T - (d - 3)\Delta t - k_1} \cdots \int_{T - (k_1 + \cdots + k_{d-2})} \frac{dk_1 \cdots dk_{d-2} dk_{d-1}}{k_1 k_2 \cdots k_{d-1}},
\]

that is,

\[
\langle F_T^{(d)} \rangle = 2 \int_{\frac{y_1}{T}}^{\frac{T}{\Delta t}} \frac{dy_1}{y_1} \int_{\frac{y_1 + \Delta t}{\Delta t}}^T \frac{dy_2}{y_2 - y_1},
\]

in which the second integral is easily computed. This leads to:

\[
\langle F_T^{(3)} \rangle = 2 \int_{\Delta t}^{T - \Delta t} \frac{dy_1}{y_1} \left[ \ln \left( \frac{T}{\Delta t} \right) + \ln \left( 1 - \frac{y_1}{T} \right) \right].
\]

(13)

Setting \( z = y_1/T \) allows one to compute the remaining integral (see also \cite{49, 50}) and yields our main result,

\[
\langle F_T^{(3)} \rangle \sim 2 \left[ \ln \left( \frac{T}{\Delta t} \right) \right]^2 + 4 \left\{ \ln \left( \frac{T}{\Delta t} \right) \ln \left( 1 - \frac{\Delta t}{T} \right) - \text{Li}_2 \left( 1 - \frac{\Delta t}{T} \right) \right\},
\]

where \( \text{Li}_2 \) is the dilogarithm function. As \( \text{Li}_2 \left( 1^- \right) \) is finite, the first term on the right hand side of Eq. (14) is also clearly the leading term when \( T/\Delta t \) becomes large. Thus,

\[
\langle F_T^{(3)} \rangle \sim 2 \left[ \ln \left( \frac{T}{\Delta t} \right) \right]^2.
\]

(14)

This asymptotical behaviour may readily be directly extracted from Eq. (13), focusing on the \( \ln \left( \frac{T}{\Delta t} \right) \) term in the integrand. This transposes straightforwardly, by simple iteration, to the \( d \)-dimensional integral in Eq (11), leading to our second main result,

\[
\langle F_T^{(d)} \rangle \sim 2 \left[ \ln \left( \frac{T}{\Delta t} \right) \right]^{d-1}.
\]

(15)

This generalizes in a very simple manner the two-dimensional result \cite{10} and the discrete-time result \cite{45}.

Both the exact formula (Eq. (14)) and the asymptotical behaviour (Eq. (15)) are confirmed in numerical simulations, as can be seen on Fig. 2. It is also the case in dimension \( d \geq 4 \).

CONCLUSION

The exact formula established here for the average number of \((d - 1)\)-dimensional facets on the convex hull of a general \(d\)-dimensional Lévy process shows how universal certain aspects of the shape of spatial random processes are. Indeed this formula is valid not only for \(d\)-dimensional Brownian motion but for a vast class of Lévy processes,
Figure 2. Numerical simulations for the average number of triangular facets on the convex hulls of some 3-dimensional Lévy processes (with $10^4$ and $10^5$ sample paths’ realizations): standard Brownian motion (pluses), and a stable process with stability index 1.5 (triangles). The solid line corresponds to the exact formula Eq. (14). The inset shows the asymptotics for large values of $T/\Delta t$, from Eq (14), with numerical simulations of 3-dimensional Brownian motion.

including all stable processes. This is reminiscent of the universality observed in the Sparre Andersen theorem [51] [52], to which the formula established here is linked via the arcsine and the uniform laws. To be more specific about the class of processes for which our formula holds: one needs the process to be truly $d$-dimensional and to have infinitely many increments on any time interval. This may not be the case for instance with compound Poisson processes. However, in many applications, one is dealing with Brownian motion or stable (self-similar) processes, which are covered here. Further work is needed to explore the cases not included, it may well be that they fall into the same universality class.

We will also seek to adapt the method introduced here to address multi-processes systems: what is the average number of facets on the global convex hull of $N$ independent spatial Brownian motions? One could also expect a large universality class in this context, and it will be of particular interest to see how the number of walkers intervenes on the $2 \ln (T/\Delta t)^{d-1}$ asymptotics in the long time limit.

---

[1] A. Einstein, Annalen der Physik 322, 549 (1905).
[2] S. Chandrasekhar, Reviews of Modern Physics 21, 383 (1949).
[3] G. Viswanathan, M. da Luz, E. Raposo, and H. Stanley, The Physics of Foraging: An Introduction to Random Searches and Biological Encounters (Cambridge University Press, 2011).
[4] O. Bénichou, M. Coppey, M. Moreau, P.-H. Suet, and R. Voituriez, Phys. Rev. Lett. 94, 198101 (2005).
[5] S. Havlin, S. Buldyrev, H. Stanley, and G. Weiss, Journal of Physics A: Mathematical and General 24, L925 (1991).
[6] S. F. Edwards, Proc. Phys. Soc. 85, 613 (1965).
[7] P. G. De Gennes, Macromolecules 9, 587 (1976).
[8] M. Westwater, Comm. Math. Phys. 72, 103 (1980).
[9] M. Doi and S. F. Edwards, The theory of polymer dynamics, International Series of Monographs on Physics (Oxford University Press, 1988).
[10] B. Worton, Ecological Modelling 38, 277 (1987).
[11] S. N. Majumdar and A. Comtet, Phys. Rev. Lett. 92, 225501 (2004).
[12] C. Haber, S. A. Ruiz, and D. Wirtz, Proceedings of the National Academy of Sciences 97, 10792 (2000).
[13] P. Cook and D. Marenduzzo, J. Cell. Biol. 188, 825 (2009).
[14] L. Takács, Am. Math. Monthly 87, 142 (1980).
[15] M. El Bachir, L’enveloppe convexe du mouvement brownien, Ph.D. thesis, Université Paul Sabatier, Toulouse, France (1983).
[16] P. Biane and G. Letac, J. Theor. Prob. 24, 330 (2011).
[17] M. Cranston, P. Hsu, and P. March, Annals Prob. 17, 144 (1989).
[18] A. Reymbaut, S. N. Majumdar, and A. Rosso, J. Phys. A: Math. Theor. 44, 415001 (2011).
