

A formula for the minimal coordination number of a parallel bundle

E. L. Starostin^{a),b)}

Centre for Nonlinear Dynamics, Department of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, London WC1E 6BT, United Kingdom

(Received 6 August 2008; accepted 8 September 2008; published online 16 October 2008)

An exact formula for the minimal coordination numbers of the parallel packed bundle of rods is presented based on an optimal thickening scenario. Hexagonal and square lattices are considered.

© 2008 American Institute of Physics. [DOI: 10.1063/1.2991413]

I. INTRODUCTION

Hexagonal packing of slender objects (filaments or rods) such as semiflexible polymers, e.g., DNA,¹⁻³ filamentous actin,⁴ collagen,⁵ or nanotubes,⁶⁻⁸ provides the most compact way of filling a domain in space. It maximizes the number of contacts between the rods. It has long been known that DNA can condensate to form bundles, in particular, closed toroids when polyvalent cations are added to the solution.⁹⁻¹³ Similar shapes are observed for other filaments stabilized by different types of adhesion: By hydrogen bonding, by ionic interactions, or by van der Waals forces.^{14,15} A perfect hexagonally packed bundle must be twistless to provide continuous parallel arrangement of filaments though it can writhe in space.¹⁶ On the other hand, the interaction peculiarities of individual filaments in a parallel bundle may favor them to fill in a nonhexagonal (e.g., square) two-dimensional (2D) lattice in cross section.

The main purpose of this paper is to present a compact exact formula for a number of unoccupied bond sites on the surface of a tightly packed parallel bundle. This quantity commonly appears in analysis of stability of DNA toroids¹⁷⁻²² and it has been computed either approximately^{17,21} or by a complicated procedure^{20,22} or for particular numbers of strands in the bundle.^{18,19} It is reasonable to believe that the exact analytical expression will prove to be handy in the future analytical research.

Consider a bundle of parallel packed N rods. In the Euclidean space, each rod may be in continuous contact with at most six other rods.²³ The total number of binding sites is $6N$ for hexagonal lattice and $4N$ for square one. Each actual bond involves two sites. Therefore, the number of the binding sites that are exposed to the environment is even and one-half of it is called the coordination number.²¹

In a poor solvent, there is a relative energetic advantage for a filament to have a polymer-polymer contact versus being exposed to solvent. Therefore, we are interested in the tightest bundle when the coordination number is minimal. The minimal coordination number enters a free energy expression that is used in studies of DNA condensation.¹⁷⁻²² Its behavior as a function of number of filaments affects stability of various conformations.

Since the rods are parallel, all orthogonal cross sections of the bundle are congruent and we can treat the problem reduced to two dimensions, i.e., we consider the set of disks that appear in an orthogonal cross section of the bundle. The minimal coordination number equals

$$\alpha_N = \lceil \sqrt{12N - 3} \rceil, \quad (1)$$

for the hexagonal lattice and

$$\beta_N = \lceil 2\sqrt{N} \rceil, \quad (2)$$

for the square lattice. In Eqs. (1) and (2), $\lceil \cdot \rceil$ stands for the ceiling function: For $x \in \mathbb{R}$, $\lceil x \rceil$ is defined as the smallest integer greater than or equal to x . In particular, $\alpha_1=3$, $\alpha_2=5$, $\alpha_3=6$, and so on and $\beta_1=2$, $\beta_2=3$, $\beta_3=4$, and so on.

Equation (1) gives exactly the same values (up to factor of 2) as a rather overcomplicated algorithm that was suggested in Ref. 20, Eq. (3), and was used recently in Ref. 22. An approximate expression was presented earlier for the average number of interactions which equals $6 - 2\alpha_N/N$ in our terms (see Eq. (10) in Ref. 17). That approximate formula may be used to compute the coordination number only for small number of members, up to $N \lesssim 30$. Actually, Eq. (10) describes a nonmonotonic function which contradicts the physical sense. Note also that the graph in Fig. 3 in the cited article which claimed to show the function from Eq. (10), in reality, represents a different nonspecified function.

For a particular case of the bundles with a complete hexagonal cross section, one can find an exact expression for the average number of interactions in Ref. 18 [Eq. (4)] which agrees with Eq. (1) if $N=3n^2+3n+1$, $n=1, 2, \dots$. It was suggested to use the formula of the number of bonds computed for complete hexagons to approximate the optimal configurations for arbitrary number of members.¹⁹ The limiting coordination number $\alpha_\infty=2\sqrt{3N}$ proposed in Ref. 21 approximates Eq. (1) for a hexagonal bundle of N filaments when $N \rightarrow \infty$. No proof was given on optimality of conformations considered in the above mentioned works to compute the minimal coordination number.

II. PROOF OF EQUATION (1)

(1) We show that Eq. (1) is valid for complete shells, i.e., for $N=M(n)=3n^2-3n+1$, $n=1, 2, \dots$ (Fig. 1). Clearly, $\alpha_1=3$ for $n=1$. For $n=2, 3, \dots$, there are six corner cells, each

^{a)}Electronic mail: e.starostin@ucl.ac.uk.

^{b)}Electronic mail: eugene_starostin@hotmail.com.

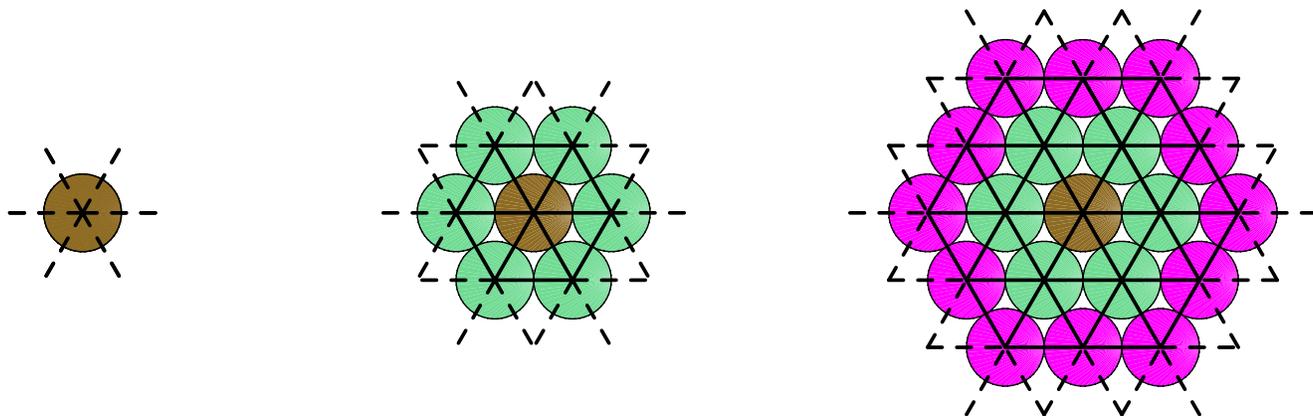


FIG. 1. (Color online) Complete shells of generation $n=1, 2, 3$. Here and in all other figures, bonds are shown as black sticks, free potential bonds are dashed.

with three open bonds and $6(n-2)$ side cells, each with two open bonds. In total, $6 \times 3 + (6n-12) \times 2 = 2(9+6n-12) = 12n-6 \equiv 2\alpha_{M(n)}$.²⁰

(2) It is easy to check the correctness of Eq. (1) for $N=1, \dots, M(2)=7$ by direct computation (Fig. 2).

Note that, as shown in Fig. 3, the configuration with minimal coordination number is obliged to be neither unique nor convex. (Convexity may be defined as absence of cells with only one free bond.)

(3) We show that Eq. (1) is valid for $N=M(n)$, $M(n)+1, \dots, M(n+1)$ for any generation number $n \geq 2$ for a chosen algorithm of adding cells.

Consider what happens when we add more cells to the complete shell of generation n ($n=3$ in Fig. 4). The first added is bound to two other cells independently of its place ($N=20$). So, it closes two existing open bonds and adds four new which means that the coordination number increases by 1: $\alpha_{M+1} = \alpha_M + \frac{1}{2}(4-2) = \alpha_M + 1$. To minimize the number of exposed sides, the next cell has to be added adjacent to the first one so that it contacts with three cells ($N=21$). Since the number of new open bonds is 3 as well, $\alpha_{M+2} = \alpha_{M+1}$.

We can keep adding new cells until the entire side of the hexagon is filled up. As it happens, the coordination number increases by 1, because the $(M(n)+n)$ -th cell can only be bound with two others ($N=22$). The cell has to be added to

one of the remaining five sides of the initial complete shell. One of the neighboring sides has to be chosen to make the next increase in the coordination number as late as possible ($N=23$). As the neighboring side is being filled, the coordination number does not change until the entire row is completed ($N=24$). The cycle repeats six times until all six sides of the $(n+1)$ -th layer are filled up (only filling two more sides is shown in Fig. 4). Each time as the $(M(n)+in)$ -th cell takes its place ($i=1, 2, 3, 4, 5$), the coordination number adds 1. There is no need for rearrangement of cells as more cells being added. Summing up, we can say that the coordination number increases each time when a cell is added to a convex configuration.

Let us now check whether Eq. (1) follows precisely this pattern. As seen from the equation, $\alpha_N = \alpha_{N-1} + \Delta$, for integers $N > 1$ and $\Delta > 0$, if and only if there exists an integer K such that $12(N-1)-3 \leq K^2 < 12N-3$. Now check what are N 's for which $\Delta > 1$. The latter would imply $\sqrt{12N-3} - \sqrt{12(N-1)-3} > 1$ which is only possible for $N=2, 3$. Since we are now interested in the case $N \geq M(2)$, we are safe to consider only $\Delta=1$. The changes in α_N may be indexed with K .

There are six changes in α_N between the shells of two consecutive generations, i.e., in the range of numbers between $M(n)$ and $M(n+1)$. Showing first that $\alpha_{M(n)+1} = \alpha_{M(n)}$

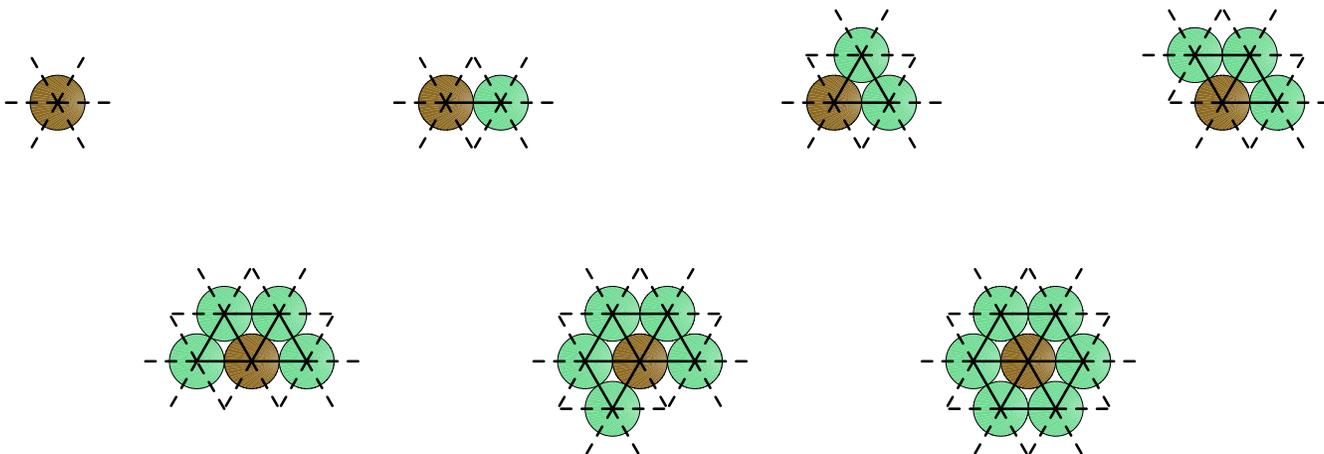


FIG. 2. (Color online) Optimal configurations for $N=M(1)=1, \dots, M(2)=7$.

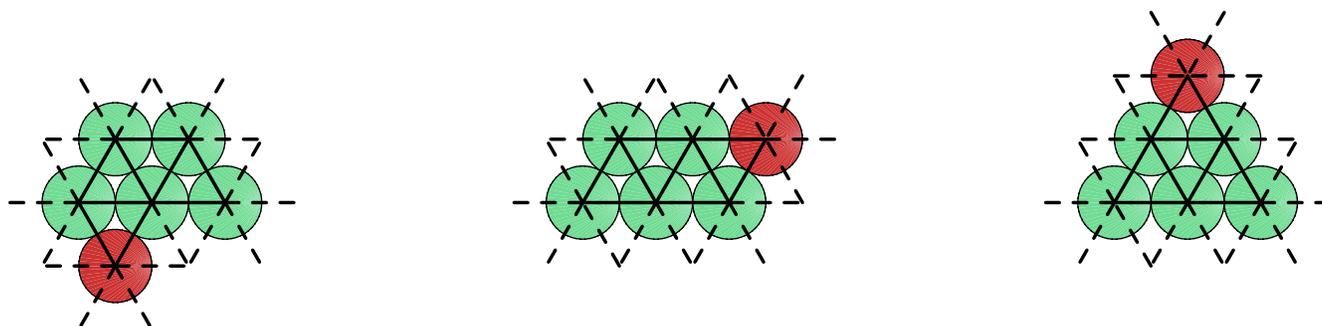


FIG. 3. (Color online) All three configurations have the minimal coordination number $\alpha_6=9$.

+1, let us set $K=\alpha_{M(n)}=6n-3$. Then $12M(n)-3 \leq (6n-3)^2 < 12(M(n)+1)-3$. Indeed, $12(3n^2-3n+1)-3=(6n-3)^2 < 12(3n^2-3n+2)-3$.

As explained above, the remaining five increases in α_N should take place for $N=M(n)+in$, $i=1,2,3,4,5$. To prove this, take $K=\alpha_{M(n)}+i$, then we can write

$$12(M(n) + in - 1) - 3 < (\alpha_{M(n)} + i)^2 < 12(M(n) + in) - 3,$$

$$12(3n^2 - 3n + 1 + in - 1) - 3 < (6n - 3 + i)^2 < 12(3n^2 - 3n + 1 + in) - 3,$$

$$36n^2 + 12(i - 3)n - 3 < 36n^2 + 12(i - 3)n + (i - 3)^2 < 36n^2 + 12(i - 3)n + 9,$$

which is obviously satisfied. Therefore, $\alpha_{M(n)+in}=\alpha_{M(n)+in-1}+1$.

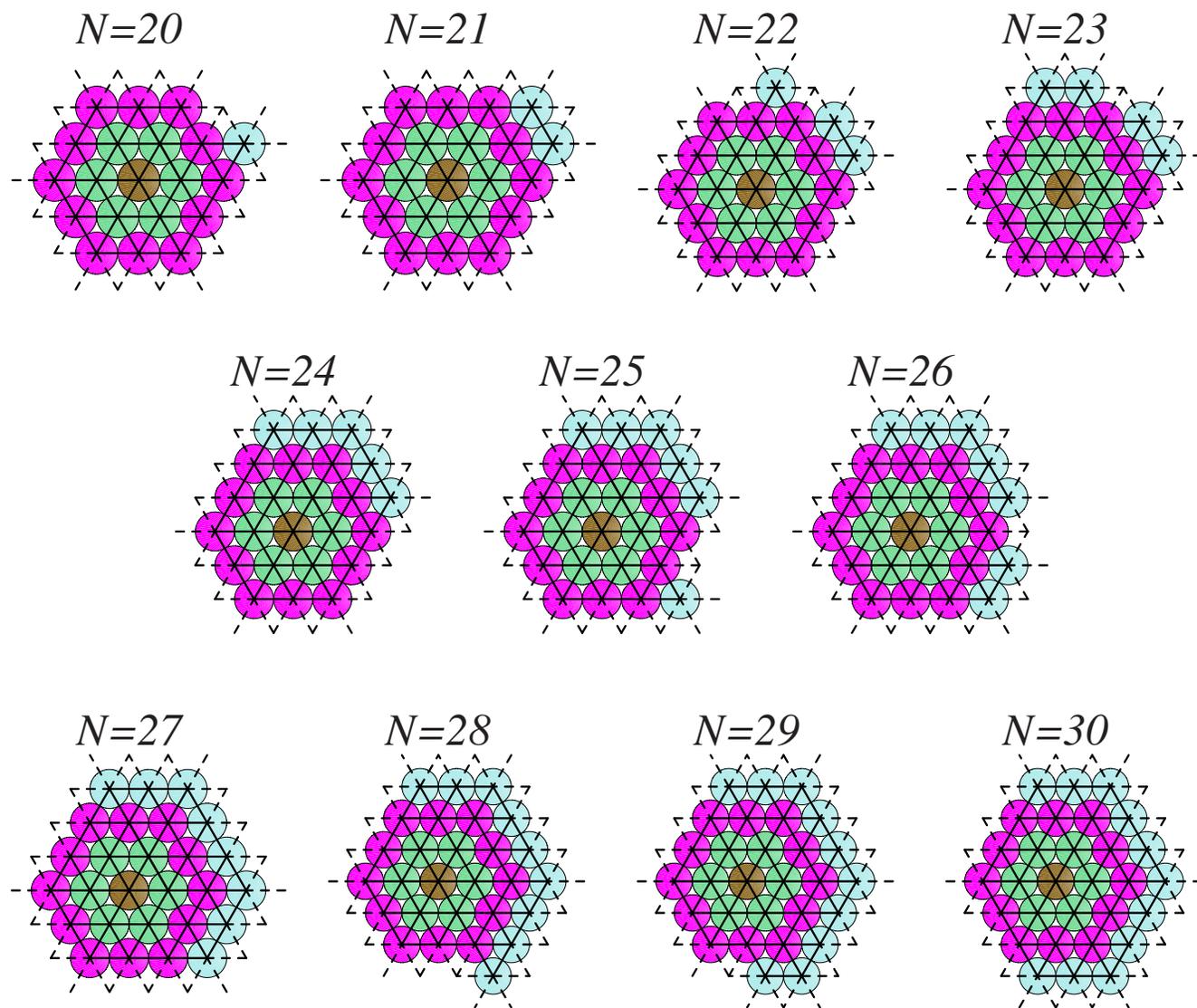


FIG. 4. (Color online) Adding more cells to the complete shell of generation $n=3$. A choice of optimal sequence for $N=20, \dots, 30$.

Thus we see that, according to our formula, α_N jumps by 1 when the first cell is added to the complete shell and when a portion of n cells is added to the complete shell of generation n unless the $n+1$ layer is completed. By way of induction we conclude that Eq. (1) is valid for all $N \geq 1$.

(4) The remaining thing to do is to show that the above scenario is optimal in the sense that it provides the minimal number of exposed sites for every number of cells. Clearly, an optimal configuration must be simply connected. Let us take an arbitrary simply connected cluster of N cells. There exists a simple closed polygonal line that connects the centers of all p cells with exposed bond sites (we call it a perimeter). Let us orient the perimeter so that the interior of the cluster ($N-p$ cells) be on the left.

Each perimeter cell can belong to only one of the following four groups. The first group consists of t cells where the perimeter line makes no turn. Such a cell has two free bonds on the right and at most two bonds with the interior part of the cluster. The second group includes u cells where the perimeter turns to the left through $\pi/3$. The members of the second group have three exposed bond sites on the right and at most one bond with the interior. The third group is made up with v cells with the $2\pi/3$ turns of the perimeter and with four free bonds on the right. The cells of this group are bonded only with two perimeter neighbors. The last fourth group is formed with w cells in which the perimeter line turns to the right through $\pi/3$. Such perimeter cells cannot appear in a convex cluster. They have one free bond on the right and up to three bonds to the interior. Thus, we have

$$p = t + u + v + w. \quad (3)$$

When traversing the complete path along the perimeter, we make exactly one turn (the winding number of the perimeter is one), hence we can write

$$u + 2v - w = 6. \quad (4)$$

By definition, the coordination number of the N -cluster α_N equals half the number of exposed bond sites on the right of the perimeter. Summing up these exposed sites gives

$$2\alpha_N = 2t + 3u + 4v + w. \quad (5)$$

Eliminating the quantities t , u , v , and w from Eq. (5) with the help of Eqs. (3) and (4) leads to

$$\alpha_N = p + 3. \quad (6)$$

The last equation simply shows that the coordination number is essentially the length of the perimeter, i.e., the number of the exposed cells.

All the bonds with the interior contribute to the coordination number of a cluster that is obtained after peeling off the N -cluster. Then, the coordination number of this ($N-p$)-cluster can be estimated as

$$2\alpha_{N-p} \leq 2t + u + 3w. \quad (7)$$

Again, with the help of Eqs. (3) and (4) we obtain from Eq. (7)

$$\alpha_{N-p} \leq p - 3.$$

Combining the last inequality with Eq. (6) allows us to relate the coordination numbers of the full and the stripped clusters

$$\alpha_N \geq \alpha_{N-p} + 6. \quad (8)$$

Let us now assume that the coordination number of the peeled-off ($N-p$)-cluster satisfies Eq. (1) for the minimal coordination number. Then we can write

$$\alpha_{N-p} \geq \sqrt{12(N-p) - 3} = \sqrt{12(N - \alpha_N) - 33}$$

[here we used Eq. (6) to eliminate p in the right-hand side]. Substitution of α_{N-p} from the last inequality into Eq. (8) and some algebraic transformations eventually lead to the required estimate for the N -cluster

$$\alpha_N \geq \sqrt{12N - 3}.$$

Thus we see that it is impossible to lower the coordination number below the value that was computed for the sequence of clusters we considered above in the former parts of this proof. We conclude that, indeed, such a scenario of growth guarantees the most compact configuration for any number of cells with the minimal coordination number from Eq. (1).

Proof of Eq. (2) is similar.

III. DISCUSSION

(1) The minimal coordination number formula may be used to compute the number of actual bonds

$$U(N) = 3N - \lceil \sqrt{12N - 3} \rceil$$

(for hexagonal lattice). For complete hexagonal shells with $N = M(n)$, $n = 1, 2, \dots$ (Fig. 1), the ceiling function can be omitted. As suggested by Ishimoto and Kikuchi,¹⁹ the expression $V(N) = 3N - \sqrt{12N - 3}$ can be used as an approximation for arbitrary N . They present a comparison of $V(N)$ for real N with what they call “the exactly counted discrete function” (though they plot two *continuous* functions). It is not clear from the paper how “the exactly counted discrete function” was computed. One can see from the graph (Fig. 2 in Ref. 19) that the latter function differs from $U(N)$ (e.g., notice too big difference between the curves for $N = 30$). Anyway, our theory provides a solid proof that $V(N)$ can be used as an approximate function for large N since it cannot differ from the exact optimal function by more than 1.

(2) Equation (1) may be “inverted” to give an equation for N_α , the maximal number of strands in the bundle for given (integer) coordination number α : $N_\alpha = \lfloor (\alpha^2 + 3)/12 \rfloor$, $\alpha \geq 5$, where $\lfloor \cdot \rfloor$ stands for the floor function: For $x \in \mathbb{R}$, $\lfloor x \rfloor$ is defined as the greatest integer less than or equal to x . For $\alpha = 9$ and $\alpha \geq 11$, the N_α -bundles are considered as particularly stable in a sense that they have the same surface energy as the $(N_\alpha - 1)$ -bundles: Addition of the N_α -th filament does not affect the coordination number. However, one more filament increases the surface energy. Such numbers N_α are called magic and supermagic.^{20,21} For the square lattice the inverted Eq. (2) reads $N_\beta = \lfloor \beta^2/4 \rfloor$, $\beta \geq 2$.

(3) A word of warning should be placed here concerning the use of the minimal coordination number to estimate the surface energy of the bundle. The coordination number, we consider in this paper, is actually defined for a thin cross-sectional slice of a bundle. To compute the surface energy of the entire bundle, one needs to integrate over its length. Since, in general, the lengths of the particular strands may differ, the total number of the exposed binding sites does not necessarily amount to simply a doubled product of the coordination number and the length of the bundle axis. To get the correct result one should account for the actual shapes of individual filaments exposed to the environment and these shapes could be rather complex. This subject is beyond the scope of this paper.

It is only worth mentioning here that, for *closed* parallel bundles the lengths of the components are quantized and this fact simplifies the computation of the surface energy reducing it essentially to the coordination number multiplied by the length. However, the life is not so simple even in this case because there are intrinsic geometric constraints on the arrangements of the rods into the bundle which are not always consistent with the minimality of the exposed surface.¹⁶ It may also be the case that the sequence of optimally packed bundles is kinetically unfeasible. For example, when the filament coils into a toroidal bundle it is unlikely that it passes through the hole.

(4) Let us focus here on the hexagonal lattice first. The complete shells are indexed with the generation number n and the number of cells is $N_\alpha = 3n^2 - 3n + 1$, $n = 1, 2, \dots$. Then, for any N -bundle we can define a real quantity $\nu_\alpha(N) \in \mathbb{R}$, $\nu_\alpha \geq 1$ such that $N = 3\nu_\alpha^2 - 3\nu_\alpha + 1$.

The problem of finding a configuration that minimizes the coordination number may be considered as an isoperimetric problem where N plays part of the “area content” and the coordination number is essentially the perimeter length [see Eq. (6)]. In the classic continuum analog, the solution of the 2D isoperimetric problem is a circle with area $A = \pi r^2$ and circumference $S = 2\pi r$ so that $(dA/dr) = S$ with the radius r being the so-called harmonic parameter.²⁴ In the discrete problem, the last equation transforms into

$$[dN/d\nu_\alpha] = [6\nu_\alpha - 3] = [\sqrt{12N - 3}] = \alpha_N.$$

Thus, for complete shells $\nu_\alpha \equiv n$ serves as the harmonic parameter and the above equation may be viewed as a discrete analog of the known expression for the perimeter of the regular polygon as a derivative of its area relative to the radius of the inscribed circle.²⁵ Note that the latter relationship follows from the Steiner–Minkowski theorem.²⁶ Thus, parallels can be seen between the sequence of the solutions to the discrete isoperimetric problem and the painting (or collaring) of polygons.

For the square lattice, we have correspondingly $\nu_\beta(N) = \sqrt{N}$ and $[dN/d\nu_\beta] = \beta_N$.

(5) The proven extremality property of layered packing

may look a little counterintuitive for thick bundles with a large number of rods. Indeed, one may expect that the optimal shape of the cross section would approach a circle as much as possible. As we can prove now, this is not the case.²¹ Moreover, the relative difference of the minimal shape and the circle becomes more noticeable as N grows. To see this let us take a complete hexagon with $M(n)$ cells. Its perimeter is a regular hexagon and it may be inscribed into a circle of radius $R = (n-1)d$, where d is the diameter of an individual rod. The maximal distance between the circle and the inscribed hexagon is proportional to the radius and equals $(1 - \sqrt{3}/2)R$. Each of six circular segments appeared between the hexagon and the circle will make a room for, say, a couple of additional layers if $n \geq 16$ ($M(16) = 721$). This suggests that these gaps could be filled in with two or more layers as more rods join the bundle. On the other hand, the optimal scenario of growth prescribes that a layer will be filled completely until the next regular hexagon with $M(n+1)$ is formed. Of course, as we have seen, the optimal configurations may not be unique (Fig. 3) and some of them could be more “circular” than others, but diversions to them in the process of growth would imply rearrangements of rods which could be prohibited.

¹D. Durand, J. Doucet, and F. Livolant, *J. Phys. II* **2**, 1769 (1992).

²F. Livolant and A. Leforestier, *Prog. Polym. Sci.* **21**, 1115 (1996).

³F. Livolant, A. M. Levelut, J. Doucet, and J. P. Benoit, *Nature (London)* **339**, 724 (1989).

⁴L. G. Tilney, *J. Cell Biol.* **64**, 289 (1975).

⁵D. J. S. Hulmes and A. Miller, *Nature (London)* **282**, 878 (1979).

⁶J.-F. Colomer, L. Henrard, G. V. Tendeloo, A. Lucas, and P. Lambin, *J. Mater. Chem.* **14**, 603 (2004).

⁷D. J. Needleman, M. A. Ojeda-Lopez, U. Raviv, H. P. Miller, L. Wilson, and C. R. Safinya, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16099 (2004).

⁸A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y. H. Lee, S. G. Kim, A. G. Rinzler, D. T. Colbert, G. E. Scuseria, D. Tománek, J. E. Fischer, and R. E. Smalley, *Science* **273**, 483 (1996).

⁹C. Böttcher, C. Endisch, J.-H. Fuhrhop, C. Catterall, and M. Eaton, *J. Am. Chem. Soc.* **120**, 12 (1998).

¹⁰Y. M. Evdokimov, A. L. Platonov, A. S. Tikhonenko, and Y. M. Varshavsky, *FEBS Lett.* **23**, 180 (1972).

¹¹R. Golan, L. I. Pietrasanta, W. Hsieh, and H. G. Hansma, *Biochemistry* **38**, 14069 (1999).

¹²N. V. Hud and I. D. Vilfan, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 295 (2005).

¹³K. A. Marx and G. C. Ruben, *Nucleic Acids Res.* **11**, 1839 (1983).

¹⁴R. Martel, H. R. Shea, and P. Avouris, *Nature (London)* **398**, 299 (1999).

¹⁵J. X. Tang, J. A. Käs, J. V. Shah, and P. A. Janmey, *Eur. Biophys. J.* **30**, 477 (2001).

¹⁶E. L. Starostin, *J. Phys.: Condens. Matter* **18**, S187 (2006).

¹⁷V. A. Bloomfield, *Biopolymers* **31**, 1471 (1991).

¹⁸A. G. Cherstvy, *J. Phys.: Condens. Matter* **17**, 1363 (2005).

¹⁹Y. Ishimoto and N. Kikuchi, *J. Chem. Phys.* **128**, 134906 (2008).

²⁰G. G. Pereira and D. R. M. Williams, *Europhys. Lett.* **50**, 559 (2000).

²¹B. Schnurr, F. Gittes, and F. C. MacKintosh, *Phys. Rev. E* **65**, 061904 (2002).

²²M. R. Stukan, E. A. An, V. A. Ivanov, and O. I. Vinogradova, *Phys. Rev. E* **73**, 051804 (2006).

²³E. L. Starostin, *Geom. Dedic.* **117**, 47 (2006).

²⁴P. Fjelstad and I. Ginchev, *Math. Mag.* **76**, 126 (2003).

²⁵J. Emert and R. Nelson, *Math. Mag.* **70**, 365 (1997).

²⁶M. Berger, *Geometry II* (Springer, New York, 1987).