Hang Lung Mathematics Awards
Vol. 10 (2023), pp. 39–94

LEAST OPTIMAL SQUARE PACKING IN A SQUARE

A RESEARCH REPORT SUBMITTED TO THE SCIENTIFIC COMMITTEE OF THE HANG LUNG MATHEMATICS AWARDS

TEAM NUMBER
2333398

TEAM MEMBERS
MOK CHUN HEI, WONG HEI

TEACHER
MR. CHEUNG TIN YAU

SCHOOL
S.K.H. TSANG SHIU TIM SECONDARY SCHOOL

AUGUST 2023

Abstract. A lot of effort has been devoted into solving the famous Square Packing Problem, which investigates the minimum side length of a square container that can pack n unit squares. This involves the search for the most optimal packing for squares. The aim of this research is to investigate an opposite idea to the original problem. We delve into the least optimal packing of squares, i.e. finding the minimum side length of a square container that can contain all configurations of n unit squares.

By considering the idea of a rotating container, we have successfully found the solution to the case of two squares. At the end, by studying the classification of intersections between the configuration and the container, as well as harnessing analytical methods, we have found the exact solution to the general case of n squares.

1. Introduction

Square packing in a square has been a famous open problem for several decades. The problem studies the optimal packing of unit squares in a larger container square such that the number of unit squares attains maximum. This problem is difficult and only a few results are found. For instance, the minimal side length of the larger container square containing 10 unit squares is $3 + \frac{1}{\sqrt{2}}$, which is one of the few known exact results. [1]

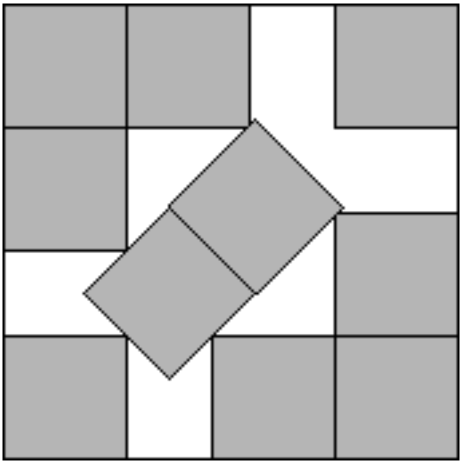


FIGURE 1. $s_{min}(10) = 3 + \frac{1}{\sqrt{2}}$.

Our investigation is induced by twisting the idea of the original problem, which seems to find the "most optimal" packing of the unit squares. Instead, we attempt to find the "least optimal" configuration of the unit squares such that the side length of its square container attains maximum. In this paper, we investigate a square container that contains all possible configurations given the number of unit squares.

Consider the following configurations of two unit squares (in gray).

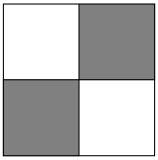


FIGURE 2. Minimum side length of the square container = 2.

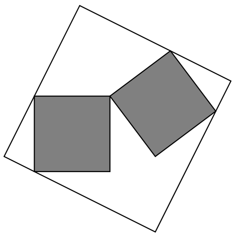


FIGURE 3. Minimum side length of the square container $= \sqrt{5}$.

Figure 3 requires a larger square container. Thus, it is a less optimal configuration compared to Figure 2. In fact, as we will see in Section 3, Figure 3 is the least optimal configuration of two unit squares. This means that any configurations of two unit squares can be contained in a square of side length $\sqrt{5}$. The goal of our research is to formulate the notion of a "least optimal" configuration, rigorously prove this fact about configurations of two unit squares, and develop tools to generalise to configurations of an arbitrary number of unit squares.

1.1. **Problem Definition.** In order to define our problem, we now define a unit square and a configuration.

Definition 1.1. *A* ***unit square*** *is a set of four points $\{A, B, C, D\}$ on a two-dimensional plane such that lengths of line segments $AB = BC = CD = AD = 1$ and $AB \perp BC$, $BC \perp CD$, $CD \perp AD$ and $AD \perp AB$. The boundary of a unit square is the union of line segments AB, BC, CD, AD. The interior of a unit square is the area bounded by the boundary excluding the boundary.*

Definition 1.2. *For a positive integer n, a set of n unit squares is called an* ***n-configuration*** *if*

- *none of the interior of the squares overlap with each other; and*
- *all the squares are connected to each other, i.e. for every pair of squares Q_1, Q_2, there exists a finite sequence of distinct squares K that starts with Q_1 and ends with Q_2 with adjacent squares in K having a point of intersection.*

In the same way the "most optimal" square packing problem requires the unit squares not to overlap, in the "least optimal" square packing problem, we restrict the unit squares to be connected since otherwise they can be brought very far away from each other and the problem is no longer interesting. In the meantime, intuitively, overlapping squares should result in a smaller square container compared to non-overlapping squares. Thus, we define a configuration that is formed by connected and also non-overlapping unit squares.

Next, we define the bounding rectangle or square, i.e. a container.

Definition 1.3. *A* ***container*** *$\mathcal{C}$ of a configuration λ is a rectangle such that all of λ lies in the interior of $\mathcal{C}$ or on the boundary of $\mathcal{C}$. Similarly, a* ***square container*** *$\mathcal{C}$ of a configuration λ is a square such that all of λ lies in the interior of $\mathcal{C}$ or on the boundary of $\mathcal{C}$.*

For the optimal square packing problem, the goal is to find the minimum length $s_{min}(n)$ such that **there exists** an n-configuration which has a square container of side length $s_{min}(n)$. Taking this as an inspiration, we define our least optimal square packing problem as follows.

Problem. *For any positive integer $n \geq 2$, find the minimum length $s(n)$ such that* ***any*** *n-configuration can be contained in a square container of side length $s(n)$.*

1.2. **Main Results.** We begin by exploring properties of the square container of configurations, which offers great insights into possible proof techniques. Then, we look into the case of 2-configuration by means of computer simulations and logical deductions, which results in $s(2) = \sqrt{5}$ as stated in Theorem 3.3. After trying several attempts and developing techniques such as classification, analysis and rearrangement of geometric figures, we lead into the complete proof, concluding that $s(n) = n+\sqrt{5}-2$. The solution is surprisingly simple compared to the thinking and proving process.

1.3. **Notation.**

Definition 1.4. *Denote* $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ *as the two-dimensional rectangular coordinate plane.*

Definition 1.5. *Denote* $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ *as the unit circle.*

Definition 1.6. *Denote* Λ_n *be the set of all* n*-configurations.*

Definition 1.7. *Denote*

$$\Lambda = \bigcup_{n=2}^{\infty} \Lambda_n$$

be the set of n*-configurations for all* n*. An element of* Λ *is known simply as a* ***configuration****.*

Definition 1.8. *For a unit vector* $\hat{r} \in \mathbb{S}^1$*,* $\hat{r}_\perp$ *is the image of* $\hat{r}$ *under anticlockwise* 90° *rotation about the origin.*

2. Square Container

2.1. **Definitions.** Let $S = \{\vec{S}_i\}$ is a finite set of points in $\mathbb{R}^2$. As a square container is essential in our research, we frequently find the length of a set of points S by projecting S along a unit vector $\hat{r}$. We denote this quantity $d_S(\hat{r})$.

Definition 2.1. *Let* $d_S^{ij}(\hat{r})$ *where* $\hat{r} \in \mathbb{S}^1$ *be the length of the line segment connecting* S_i *and* S_j *when projected along* $\hat{r}$ *i.e.*

$$d_S^{ij}(\hat{r}) = |\hat{r} \cdot (\vec{S}_i - \vec{S}_j)|.$$

Definition 2.2. *Let* $d_S(\hat{r})$ *where* $\hat{r} \in \mathbb{S}^1$ *be*

$$d_S(\hat{r}) = \max\{\hat{r} \cdot \vec{P} : \vec{P} \in S\} - \min\{\hat{r} \cdot \vec{P} : \vec{P} \in S\} = \max_{i,j} d_S^{ij}(\hat{r}).$$

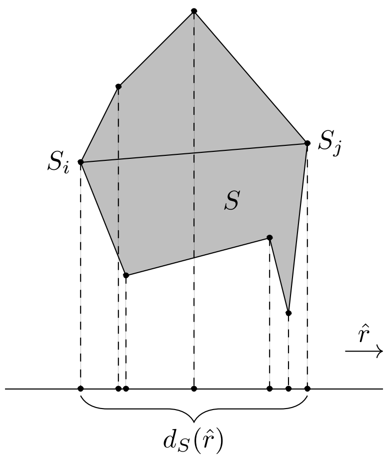


FIGURE 4. Illustration of $d_S(\hat{r})$.

As a result, for a square container along the directions $\hat{r}$ and $\hat{r}_\perp$, the minimum side length is the greater of $d_S(\hat{r})$ and $d_S(\hat{r}_\perp)$, denoted $f_S(\hat{r})$.

Definition 2.3. *Let the side length of the square container of S when oriented along unit vector $\hat{r}$ be*

$$f_S(\hat{r}) = \max\{d_S(\hat{r}), d_S(\hat{r}_\perp)\}.$$

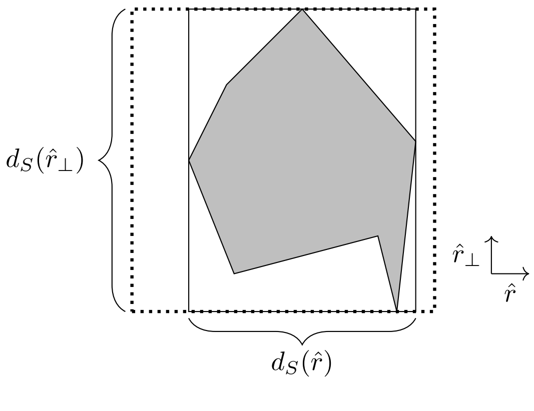


FIGURE 5. Illustration of $f_S(\hat{r})$. The smallest square container along directions $\hat{r}$ and $\hat{r}_\perp$, shown in dotted line, has side length $\max\{d_S(\hat{r}), d_S(\hat{r}_\perp)\}$.

Definition 2.4. *The side length of the smallest square container of S is the minimum value of f_S, denoted f_S^{min}:*

$$f_S^{min} = \min\{f_S(\hat{r}) : \hat{r} \in \mathbb{S}^1\}.$$

Note that each configuration is a set of unit squares, which by definition is a set of points. Let us define the notation of d_λ, f_λ etc. where λ is a configuration.

Definition 2.5. *For a configuration λ, let T be the set of all vertices of all unit squares in λ. Then $d_\lambda^{ij}(\hat{r})$, $d_\lambda(\hat{r})$, $f_\lambda(\hat{r})$ and f_λ^{min} are defined the same as $d_T^{ij}(\hat{r})$, $d_T(\hat{r})$, $f_T(\hat{r})$ and f_T^{min} respectively.*

Then our problem can be rephrased as: for a positive integer $n \geq 2$, $s(n)$ is defined as the maximum side length of the smallest square container of any n-configuration:

$$s(n) = \max\{f_\lambda^{min} : \lambda \in \Lambda_n\}.$$

2.2. **Properties.** We will now discuss some of the properties of a square container of a set of points S.

First, the following two lemmas describe basic properties of the functions d_S and f_S which are trivial to see.

Lemma 2.6. $d_S(\hat{r}) = d_S(-\hat{r})$.

Lemma 2.7. $f_S(\hat{r}) = f_S(\hat{r}_\perp) = f_S(-\hat{r})$.

Next, we will show that there always exists an orientation such that a rectangle container oriented that way is a square.

Lemma 2.8. *For a finite set of points S, there exists some $\hat{r}_0 \in \mathbb{S}^1$ such that*

$$d_S(\hat{r}_0) = d_S(\hat{r}_{0\perp}).$$

Proof. Let $g(\hat{r}) = d_S(\hat{r}) - d_S(\hat{r}_\perp)$. We will show that $g(\hat{r}_0) = 0$ for some $\hat{r}_0 \in \mathbb{S}^1$.

Pick any $\hat{r}' \in \mathbb{S}^1$. If $g(\hat{r}') = 0$, we are done. Otherwise notice that

$$g(\hat{r}'_\perp) = d_S(\hat{r}'_\perp) - d_S(\hat{r}'_{\perp_\perp}) = d_S(\hat{r}'_\perp) - d_S(\hat{r}') = -g(\hat{r}').$$

Since d_S is continuous, so is g. Therefore, by the Intermediate Value Theorem, there exists $\hat{r}_0 \in \mathbb{S}^1$ such that $g(\hat{r}_0) = 0$. □

This turns out to be useful when simplifying an n-configuration, which we will use in Lemma 5.2.

Moreover, one of our goals is to find the value of f_S^{min}, which is the minimum value of the maximum of a collection of functions. So, it is helpful to probe into the behaviours of this function. Using analytical methods, it is possible to categorise a local minimum into cases, specified in Lemma 2.9.

Lemma 2.9. *For $h(x) = \max\{f(x), g(x)\}$ where $f(x)$ and $g(x)$ are real-valued continuous functions, h is at a local minimum at $x = x_0$ only if one of the following is true:*

(a) *f is at a local minimum at $x = x_0$;*
(b) *g is at a local minimum at $x = x_0$; or*
(c) *$f(x_0) = g(x_0)$.*

Proof. Let $y = h(x_0)$. Then this happens only if one of the following is true:

(1) $g(x_0) < f(x_0) = y$;
(2) $f(x_0) < g(x_0) = y$;

(3) $f(x_0) = g(x_0) = y$, which corresponds to case (c).

We focus on case 1. Choose a value $z \in (g(x_0), f(x_0))$. By continuity of f, there exists $\delta_f > 0$ such that for all $x \in (x_0 - \delta_f, x_0 + \delta_f)$, $f(x) > z$. Then by continuity of g, there exists $\delta_g > 0$ such that for all $x \in (x_0 - \delta_g, x_0 + \delta_g)$, $g(x) < z$.

Set $\delta_h = \min\{\delta_f, \delta_g\} > 0$. Then for all $x \in (x_0 - \delta_h, x_0 + \delta_h)$, $f(x) > z > g(x)$, so $h(x) = f(x)$. Therefore by considering the neighbourhood $(x_0 - \delta_h, x_0 + \delta_h)$ around x_0, h is at a local minimum only if f is at a local minimum, which proves case (a).

Case 2 can be proven similarly to be equivalent to case (b). □

This can then be generalised to the maximum of any finite number of functions using induction.

Corollary 2.10. *For $h(x) = \max\{f_i(x)\}$ where $i = 1, 2, \ldots, n$ and $f_i(x)$ are all real-valued continuous functions, h is at a local minimum at $x = x_0$ only if one of the following is true:*

(a) *f_i is at a local minimum at $x = x_0$ for some $i = 1, 2, \ldots, n$; or*
(b) *$f_i(x_0) = f_j(x_0)$ for some distinct $i, j = 1, 2, \ldots, n$.*

Proof. We will prove this by induction. For $n = 2$, this is proven in Lemma 2.9.

Assume this is true for $n = k$. For $n = k + 1$, $h(x) = \max\{g(x), f_{k+1}(x)\}$ where $g(x) = \max\{f_1(x), \ldots, f_k(x)\}$. By Lemma 2.9, h is at a local minimum only if one of the following is true:

(1) f_{k+1} is at a local minimum;
(2) g is at a local minimum; or
(3) $f_{k+1}(x_0) = g(x_0)$.

Induction assumption gives that case 2 happens only if

- f_i is at a local minimum for some $i = 1, 2, \ldots, k$; or
- $f_i(x_0) = f_j(x_0)$ for some distinct $i, j = 1, 2, \ldots, k$.

And case 3 happens only if $f_{k+1}(x_0) = f_i(x_0)$ for some $i = 1, 2, \ldots, k$.

Rewriting the cases,

- f_{k+1} is at a local minimum;
- f_i is at a local minimum for some $i = 1, 2, \ldots, k$;
- $f_i(x_0) = f_j(x_0)$ for some distinct $i, j = 1, 2, \ldots, k$; or
- $f_{k+1}(x_0) = f_i(x_0)$ for some $i = 1, 2, \ldots, k$,

hence the case when $n = k + 1$ is proven. □

We can apply this characteristic of local minima to d_S and f_S as follows.

Lemma 2.11. *If $d_S(\hat{r}) \neq 0$ for all $\hat{r} \in \mathbb{S}^1$, $d_S(\hat{r})$ is at a local minimum at $\hat{r} = \hat{r}_0$ only if*

$$d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_0)$$

for some i, j, k, l where $i < j$, $k < l$, and $(i, j) \neq (k, l)$.

Proof. Note that for all i, j and $\hat{r}$, $d_S^{ii}(\hat{r}_0) = 0$ and also $d_S^{ij}(\hat{r}) = d_S^{ji}(\hat{r})$. Given that $d_S(\hat{r}) \neq 0$, d_S can be simplified to

$$d_S(\hat{r}) = \max_{i<j} d_S^{ij}(\hat{r}).$$

By Corollary 2.10, d_S is at a local minimum only if

(1) $d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_0)$ for some i, j, k, l where $i < j, k < l, (i, j) \neq (k, l)$; or
(2) d_S^{ij} is at a local minimum for some i, j.

Note that the local minimum of $d_S^{ij}(\hat{r}) = |\hat{r} \cdot (\vec{S}_i - \vec{S}_j)|$ is 0 for all i, j, so case 2 is impossible considering $d_S \neq 0$. Thus $d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_0)$ for some i, j, k, l where $i < j, k < l, (i, j) \neq (k, l)$. □

Theorem 2.12. *If $d_S(\hat{r}) \neq 0$ for all $\hat{r} \in \mathbb{S}^1$, f_S is at a local minimum at $\hat{r} = \hat{r}_0$ only if for some i, j, k, l with $i < j$ and $k < l$, one of the following is true:*

(a) $d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_0)$ *where* $(i, j) \neq (k, l)$;
(b) $d_S^{ij}(\hat{r}_{0\perp}) = d_S^{kl}(\hat{r}_{0\perp})$ *where* $(i, j) \neq (k, l)$; *or*
(c) $d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_{0\perp})$.

Proof. Using the definition of f_S and Lemma 2.9, f_S attains local minimum only if one of the following is true:

(1) $d_S(\hat{r})$ is at a local minimum;
(2) $d_S(\hat{r}_\perp)$ is at a local minimum; or
(3) $d_S(\hat{r}) = d_S(\hat{r}_\perp)$.

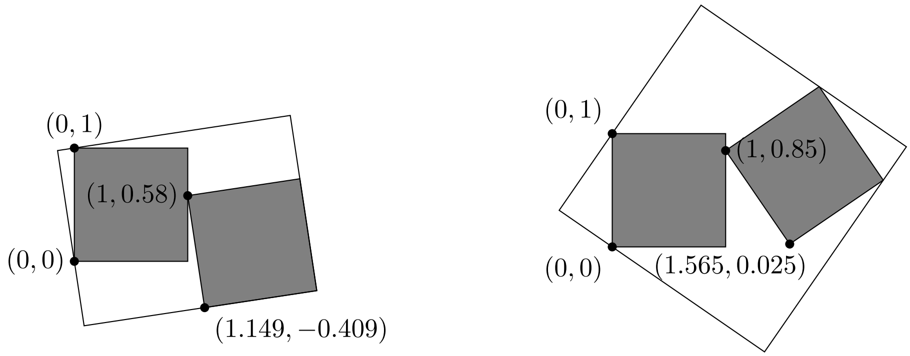
Case 1 occurs, by Lemma 2.11, only if $d_S^{ij}(\hat{r}_0) = d_S^{kl}(\hat{r}_0)$ for some i, j, k, l with $i < j$, $k < l$, $(i, j) \neq (k, l)$, which corresponds to case (a).

Case 2 occurs similarly only if $d_S^{ij}(\hat{r}_{0\perp}) = d_S^{kl}(\hat{r}_{0\perp})$ for some i, j, k, l with $i < j$, $k < l$, $(i, j) \neq (k, l)$, which corresponds to case (b).

For case 3, using definition of d_S, $d_S(\hat{r}_0) = d_S^{ij}(\hat{r})$ and $d_S(\hat{r}_{0\perp}) = d_S^{kl}(\hat{r}_{0\perp})$ for some i, j, k, l. $i \neq j$ and $k \neq l$ since $d_S(\hat{r}_0)$ and $d_S(\hat{r}_{0\perp})$ are non-zero. This corresponds to case (c). □

This gives an important insight in finding f_S^{min}. If f_S attains a minimum at $\hat{r}$ and we have a rectangle $\mathcal{C}$ oriented along direction $\hat{r}$ and $\hat{r}_\perp$ that tightly contains S, either $\mathcal{C}$ is a square (which we call an α container), or one of its side length, i.e. $d_S(\hat{r})$ or $d_S(\hat{r}_\perp)$, attains a local minimum (which we call a β container).

These properties give us a clearer picture of the behaviour of f_S^{min}. Combining with the necessary existence of an α container proven in Lemma 2.8, although Theorem 2.12 does not contribute logically to our final results, it guides us towards considering the α container as a key to this problem. Besides, our experience in finding the numerical value of f_S^{min} in graphing software reveals that the β case usually results in f_S^{min} that is far smaller than $s(2)$, while the α case usually results in f_S^{min} that is close to $s(2)$. Figure 6 gives an example of these two cases. As a result, these observations give us strong motivation to focus on α containers of configurations.



(A) The β container that corresponds to f_S^{min}. $f_S^{min} \approx 2.075$, which is far smaller than $s(2)$.

(B) The α container that corresponds to f_S^{min}. $f_S^{min} \approx 2.201$, which is close to $s(2)$.

FIGURE 6. An α and a β container of a configuration.

Finally, to facilitate focusing on the α case, we prove the following lemma to find an upper bound of $s(n)$.

Lemma 2.13. *Let $\hat{r}_c : \Lambda_n \to \mathbb{S}^1$ be a function such that for each n-configuration λ, choose a unit vector $\hat{r}_c(\lambda)$. Then the maximum of $f_\lambda\big(\hat{r}_c(\lambda)\big)$ for all $\lambda \in \Lambda_n$ is an upper bound for $s(n)$:*

$$s(n) \leq \max\{f_\lambda\big(\hat{r}_c(\lambda)\big) : \lambda \in \Lambda_n\}.$$

Proof. For all $\lambda' \in \Lambda_n$, by the definition of $f_{\lambda'}^{min}$,

$$f_{\lambda'}^{min} = \min\{f_{\lambda'}(\hat{r}) : \hat{r} \in \mathbb{S}^1\} \leq f_{\lambda'}\big(\hat{r}_c(\lambda')\big) \leq \max\{f_\lambda\big(\hat{r}_c(\lambda)\big) : \lambda \in \Lambda_n\}$$

so

$$s(n) = \max\{f_\lambda^{min} : \lambda \in \Lambda_n\} \leq \max\{f_\lambda\big(\hat{r}_c(\lambda)\big) : \lambda \in \Lambda_n\}.$$

□

This lemma allows us to not consider f_λ^{min}. Instead, we can always choose a specific orientation of the square container to calculate an upper bound which is sufficient to prove our result. In our later proof for $s(n)$, we choose an orientation that aims to make the container an α case.

3. 2-CONFIGURATION

We construct a computer program to iterate through y (as `y` in the code) and θ (as `t` in the code) to generate all possible 2-configurations. It also iterates through the angle of the container (as `angle_of_r` in the code), hence a total of 3 layers of nested iteration i.e. a time complexity of $O(n^3)$. Refer to the appendix for the C++ code and Figure 12 for the notations.

The computer program gives an approximate answer of $s(2) \approx 2.236515$, which is close to $\sqrt{5}$. With the aid of graphing software, we have also confirmed that $s(2)$ is close to $\sqrt{5}$. Thus, it is very likely that the value of $s(2)$ is $\sqrt{5}$.

We will find the value of $s(2)$ in two steps:

(1) proving that $\sqrt{5}$ is a lower bound of $s(2)$ by showing there exists a 2-configuration with $f_\lambda^{min} = \sqrt{5}$; then
(2) proving that $\sqrt{5}$ is an upper bound of $s(2)$ by showing all 2-configurations can be contained in a square of side length $\sqrt{5}$.

Lemma 3.1. $s(2) \geq \sqrt{5}$.

Proof. We will show that there exists $\lambda \in \Lambda_2$ such that $f_\lambda^{min} = \sqrt{5}$. Consider the following points:

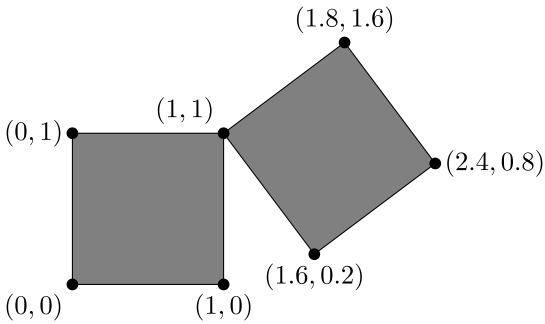


FIGURE 7. "Least optimal" 2-configuration.

By considering the points i, j where $d_\lambda(\hat{r}) = d_\lambda^{ij}(\hat{r})$ for different ranges of θ, $d_\lambda(\hat{r})$ can be written as the following partial function.

$$d_\lambda((\cos\theta, \sin\theta)) = \begin{cases} \frac{4\sqrt{10}}{5}\cos\left(\theta - \tan^{-1}\frac{1}{3}\right), & \theta \in [0, \tan^{-1}\frac{3}{4}) \\ \frac{\sqrt{145}}{5}\cos\left(\theta - \tan^{-1}\frac{8}{9}\right), & \theta \in [\tan^{-1}\frac{3}{4}, \frac{\pi}{2}) \\ \frac{4\sqrt{5}}{5}\cos\left(\theta - \tan^{-1}2\right), & \theta \in [\frac{\pi}{2}, \pi - \tan^{-1}3) \\ -\frac{4\sqrt{5}}{5}\cos\left(\theta + \tan^{-1}\frac{1}{2}\right), & \theta \in [\pi - \tan^{-1}3, \pi - \tan^{-1}\frac{4}{3}) \\ -\frac{\sqrt{145}}{5}\cos\left(\theta + \tan^{-1}\frac{1}{12}\right), & \theta \in [\pi - \tan^{-1}\frac{4}{3}, \pi) \end{cases}$$

Figures 8 to 11 show the rotating container. The black dots are points i, j, which determine the side length of the container of this configuration for certain orientations.

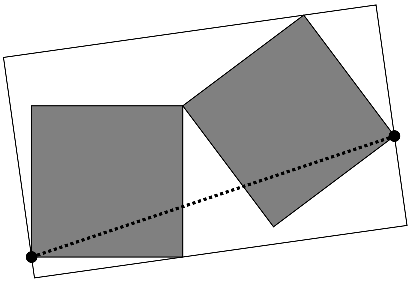


FIGURE 8. The container for $\theta \in \big[0, \tan^{-1}\frac{3}{4}\big)$.

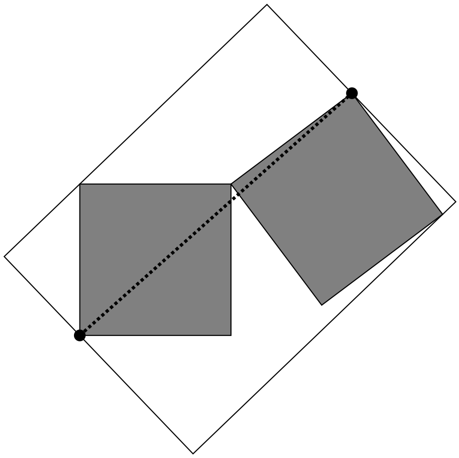


FIGURE 9. The container for $\theta \in \big[\tan^{-1}\frac{3}{4}, \tan^{-1}2\big)$.

Note that $d_\lambda(\hat{r}) < \sqrt{5}$ for all $\theta \in (\tan^{-1}2, \frac{\pi}{2} + \tan^{-1}2)$ while $d_\lambda(\hat{r}) \geq \sqrt{5}$ for all $\theta \in [0, \tan^{-1}2] \cup [\frac{\pi}{2} + \tan^{-1}2, \pi)$. Thus,

$$f_\lambda((\cos\theta, \sin\theta)) = \max\Big\{d_\lambda((\cos\theta, \sin\theta)), d_\lambda\Big(\big(\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})\big)\Big)\Big\}$$

$$= \begin{cases} \frac{4\sqrt{10}}{5}\cos\big(\theta - \tan^{-1}\frac{1}{3}\big), & \theta \in [0, \tan^{-1}\frac{3}{4}) \\ \frac{\sqrt{145}}{5}\cos\big(\theta - \tan^{-1}\frac{8}{9}\big), & \theta \in [\tan^{-1}\frac{3}{4}, \tan^{-1}2) \\ \frac{\sqrt{145}}{5}\cos\big(\theta - \tan^{-1}12\big), & \theta \in [\tan^{-1}2, \frac{\pi}{2}) \\ -\frac{4\sqrt{10}}{5}\cos\big(\theta + \tan^{-1}3\big), & \theta \in [\frac{\pi}{2}, \pi - \tan^{-1}\frac{4}{3}) \\ -\frac{\sqrt{145}}{5}\cos\big(\theta + \tan^{-1}\frac{9}{8}\big), & \theta \in [\pi - \tan^{-1}\frac{4}{3}, \pi - \tan^{-1}\frac{1}{2}) \\ -\frac{\sqrt{145}}{5}\cos\big(\theta + \tan^{-1}\frac{1}{12}\big), & \theta \in [\pi - \tan^{-1}\frac{1}{2}, \pi) \end{cases}$$

which attains the minimum $\sqrt{5}$ at $\theta = \tan^{-1}2$ or $\frac{\pi}{2} + \tan^{-1}2$. □

The following lemma describes an algebraic fact that is used in Theorem 3.3.

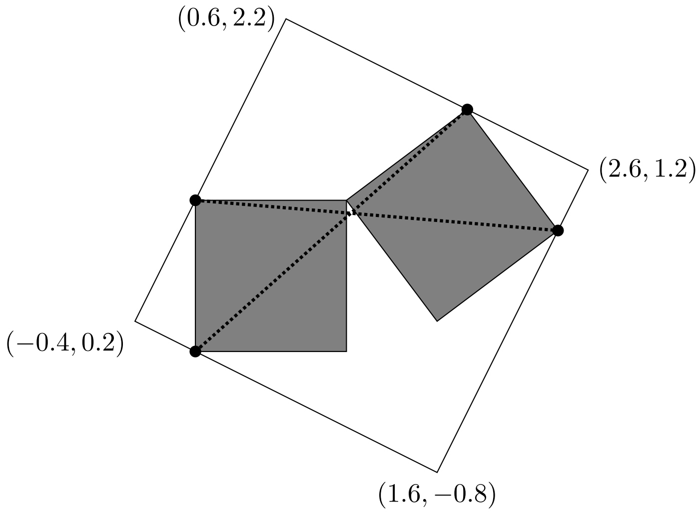


FIGURE 10. The container at $\theta = \tan^{-1} 2$.

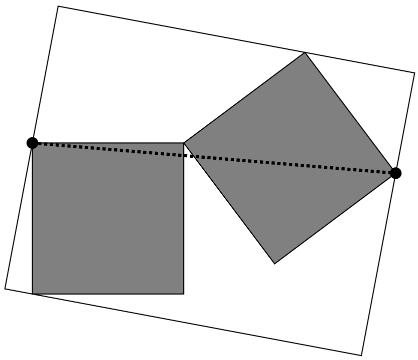


FIGURE 11. The container for $\theta \in \left[\tan^{-1} 2, \frac{\pi}{2}\right)$.

Lemma 3.2. *For a real-valued quadratic function $f(x) = ax^2 + bx + c$ with $a > 0$ which has real roots, if a real number $p \leq -\frac{b}{2a}$ satisfies $f(p) \geq 0$, then p is less than or equal to any roots of $f(x)$.*

Proof. Let α and β be the roots of $f(x)$ where $\alpha \leq \beta$. Then

x	$(-\infty, \alpha)$	(α, β)	(β, ∞)
$f(x)$	$+$	$-$	$+$

so $p \leq \alpha$ or $p \geq \beta$.
Also

$$p \leq -\frac{b}{2a} = \frac{\alpha + \beta}{2} \leq \frac{\beta + \beta}{2} = \beta$$

with the equality occurring only when $\alpha = \beta$. Therefore $p \leq \alpha \leq \beta$. □

We can now prove the upper bound of $s(2)$ and therefore its exact value. This proof mainly uses methods from algebra and calculus.

Theorem 3.3. $s(2) = \sqrt{5}$.

Proof. We will show that $s(2) \leq \sqrt{5}$ using Lemma 2.13, and combining with Lemma 3.1 proves $s(2) = \sqrt{5}$.

Any configuration $\lambda \in \Lambda_2$ can be transformed into Figure 12 under translation, rotation, reflection for some $y \in [0, 1]$ and $\theta \in [0, \frac{\pi}{2}]$.

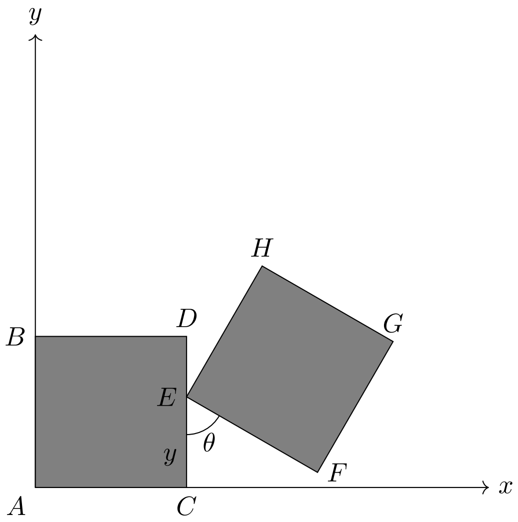


FIGURE 12

The set of points is $S = \{A, B, C, D, E, F, G, H\}$ where

- $A = (0, 0)$
- $B = (0, 1)$
- $C = (1, 0)$
- $D = (1, 1)$
- $E = (1, y)$
- $F = (1 + \sin\theta, y - \cos\theta)$
- $G = (1 + \sin\theta + \cos\theta, y + \sin\theta - \cos\theta)$
- $H = (1 + \cos\theta, y + \sin\theta)$

[3] gave a method to, given four points on a plane, construct a square such that each side of the square (or its extension) would pass through exactly one of the given points. Namely, given 4 points A, B, C, D (assume A and C lie on the opposite sides of the future square, same for B and D), perform the following steps.

(1) Join AC.
(2) Drop a perpendicular line from B to AC and find E on that perpendicular line such that $BE = AC$.
(3) D and E lie on the same side of the square (or its extension).

(4) The rest of the construction consists of dropping perpendicular lines to form the other three sides.

Furthermore, the existence of such square is proved by Lemma 2.8.

We now choose an orientation such that A, B, G, H are on each side of a square of that orientation. Using the construction method, let a point $I = (y+\sin\theta, -\cos\theta)$. Then $AH = BI$ and $AH \perp BI$. Let $\hat{r}$ be the unit vector along $\overrightarrow{IG}$:

$$\hat{r} = \frac{1}{k}(1 - y + \cos\theta, y + \sin\theta)$$

where $k = \sqrt{(1 - y + \cos\theta)^2 + (y + \sin\theta)^2}$. Then

$$\hat{r}_\perp = \frac{1}{k}(-y - \sin\theta, 1 - y + \cos\theta).$$

To maximise $d_\lambda(\hat{r})$ and $d_\lambda(\hat{r}_\perp)$, we consider the dot products of each of the position vectors of the points with $\hat{r}$ and $\hat{r}_\perp$:

P	$k\hat{r} \cdot \vec{P}$	$k\hat{r}_\perp \cdot \vec{P}$
A	0	0
B	$y + \sin\theta$	$1 - y + \cos\theta$
C	$1 - y + \cos\theta$	$-y - \sin\theta$
D	$1 + \sin\theta + \cos\theta$	$1 - 2y - \sin\theta + \cos\theta$
E	$1 + \cos\theta - y(1 + \sin\theta) + y^2$	$-\sin\theta + y\cos\theta - y^2$
F	$1+\sin\theta+\cos\theta-y(1+\cos\theta)+y^2$	$-1 - \sin\theta - \cos\theta - y(\sin\theta - 2\cos\theta) - y^2$
G	$2 + \sin\theta + 2\cos\theta + y(\sin\theta - 2\cos\theta - 1) + y^2$	$-1-\cos\theta+y(\cos\theta-2\sin\theta)-y^2$
H	$2+2\cos\theta+y(2\sin\theta-\cos\theta-1)+y^2$	$-y\sin\theta - y^2$

TABLE 1. $k\hat{r} \cdot \vec{P}$ and $k\hat{r}_\perp \cdot \vec{P}$.

Therefore,

$$\begin{aligned} d_\lambda(\hat{r}) &= \max\{\hat{r} \cdot \vec{P} : \vec{P} \in S\} - \min\{\hat{r} \cdot \vec{P} : \vec{P} \in S\} \\ &= \max\{\hat{r} \cdot \vec{G}, \hat{r} \cdot \vec{H}\} - \hat{r} \cdot \vec{A} \\ &= \max\{\hat{r} \cdot (\vec{G} - \vec{A}), \hat{r} \cdot (\vec{H} - \vec{A})\} \\ &= \max\left\{\frac{2 + \sin\theta + 2\cos\theta + y(\sin\theta - 2\cos\theta - 1) + y^2}{\sqrt{(1 - y + \cos\theta)^2 + (y + \sin\theta)^2}},\right. \\ &\quad \left.\frac{2 + 2\cos\theta + y(2\sin\theta - \cos\theta - 1) + y^2}{\sqrt{(1 - y + \cos\theta)^2 + (y + \sin\theta)^2}}\right\} \end{aligned}$$

and

$$\begin{aligned} d_\lambda(\hat{r}_\perp) &= \max\{\hat{r}_\perp \cdot \vec{P} : \vec{P} \in S\} - \min\{\hat{r}_\perp \cdot \vec{P} : \vec{P} \in S\} \\ &= \hat{r}_\perp \cdot \vec{B} - \min\{\hat{r} \cdot \vec{F}, \hat{r} \cdot \vec{G}\} \\ &= \max\{\hat{r}_\perp \cdot (\vec{B} - \vec{F}), \hat{r}_\perp \cdot (\vec{B} - \vec{G})\}. \end{aligned}$$

Notice that $\hat{r}_\perp \cdot (\vec{B} - \vec{F}) = \hat{r} \cdot (\vec{G} - \vec{A})$ and $\hat{r}_\perp \cdot (\vec{B} - \vec{G}) = \hat{r} \cdot (\vec{H} - \vec{A})$, so $d_\lambda(\hat{r}) = d_\lambda(\hat{r}_\perp)$. Thus

$$f_\lambda(\hat{r}) = d_\lambda(\hat{r}).$$

Also, the two expressions being maximised in $d_\lambda(\hat{r})$ are equivalent after a substitution of $\theta \mapsto \pi/2 - \theta$ and $y \mapsto 1 - y$:

$$\begin{aligned} &\frac{2 + \sin(\frac{\pi}{2} - \theta) + 2\cos(\frac{\pi}{2} - \theta) + (1-y)(\sin(\frac{\pi}{2} - \theta) - 2\cos(\frac{\pi}{2} - \theta) - 1) + (1-y)^2}{\sqrt{(1 - (1-y) + \cos(\frac{\pi}{2} - \theta))^2 + (1 - y + \sin(\frac{\pi}{2} - \theta))^2}} \\ &= \frac{2 + 2\cos\theta + y(2\sin\theta - \cos\theta - 1) + y^2}{\sqrt{(1 - y + \cos\theta)^2 + (y + \sin\theta)^2}}. \end{aligned}$$

Hence, when we take the maximum of $d_\lambda(\hat{r})$ for all values of y and θ, we can disregard one of the two expressions.

$$\begin{aligned} &\max\{f_\lambda(\hat{r}) : \lambda \in \Lambda_2\} \\ &= \max\{d_\lambda(\hat{r}) : \lambda \in \Lambda_2\} \\ &= \max\left\{\frac{2 + 2\cos\theta + y(2\sin\theta - \cos\theta - 1) + y^2}{\sqrt{(1 - y + \cos\theta)^2 + (y + \sin\theta)^2}} : y \in [0,1], \theta \in \left[0, \frac{\pi}{2}\right]\right\} \\ &= \max\left\{\frac{2 + 2\cos\theta + y(2\sin\theta - \cos\theta - 1) + y^2}{\sqrt{2 + 2\cos\theta + y(2\sin\theta - 2\cos\theta - 2) + 2y^2}} : y \in [0,1], \theta \in \left[0, \frac{\pi}{2}\right]\right\} \\ &= \max\left\{\frac{\alpha + \beta y + y^2}{\sqrt{\alpha + \gamma y + 2y^2}} : y \in [0,1], \theta \in \left[0, \frac{\pi}{2}\right]\right\} \end{aligned}$$

where $\alpha = 2 + 2\cos\theta, \beta = 2\sin\theta - \cos\theta - 1, \gamma = 2\sin\theta - 2\cos\theta - 2$.

To maximise $d_\lambda(\hat{r})$, we take its partial derivative with respect to y,

$$\begin{aligned} \frac{\partial d_\lambda(\hat{r})}{\partial y} &= \frac{(\alpha + \gamma y + 2y^2)(\beta + 2y) - \frac{1}{2}(\alpha + \beta y + y^2)(\gamma + 4y)}{(\alpha + \gamma y + 2y^2)^{\frac{3}{2}}} \\ &= \frac{\alpha(\beta - \frac{1}{2}\gamma) + \frac{1}{2}\beta\gamma y + \frac{3}{2}\gamma y^2 + 2y^3}{(\alpha + \gamma y + 2y^2)^{\frac{3}{2}}} \\ &= \frac{(y + \sin\theta)(2y^2 + (\sin\theta - 3\cos\theta - 3)y + (2 + 2\cos\theta))}{(\alpha + \gamma y + 2y^2)^{\frac{3}{2}}}. \end{aligned}$$

We will show that $\frac{\partial d_\lambda(\hat{r})}{\partial y} \geq 0$ for all $y \in [0,1]$. Consider the expression

$$2y^2 + (\sin\theta - 3\cos\theta - 3)y + (2 + 2\cos\theta). \tag{1}$$

If Expression (1) has roots, then its discriminant satisfies

$$\begin{aligned}0 &\leq (\sin\theta - 3\cos\theta - 3)^2 - 4(2)(2 + 2\cos\theta)\\ &= \sin^2\theta + 9\cos^2\theta + 9 - 6\sin\theta\cos\theta + 18\cos\theta - 6\sin\theta - 16 - 16\cos\theta\\ &= 8\cos^2\theta + 2\cos\theta - 6\sin\theta\cos\theta - 6\sin\theta - 6\\ &= 2(1+\cos\theta)(4\cos\theta - 3\sin\theta - 3)\end{aligned}$$

so

$$4\cos\theta - 3\sin\theta - 3 \geq 0.$$

We now verifies the conditions in Lemma 3.2 to show that all roots of Expression (1) is larger than or equal to 1. Since when substituting $y = 1$,

$$2(1)^2 + (\sin\theta - 3\cos\theta - 3)(1) + (2 + 2\cos\theta) = 1 - \cos\theta + \sin\theta \geq 0$$

and its mean of roots

$$\begin{aligned}-\frac{\sin\theta - 3\cos\theta - 3}{2(2)} &= \frac{1}{4}((4\cos\theta - 3\sin\theta - 3) + 6 + 2\sin\theta - \cos\theta)\\ &\geq \frac{1}{4}(0 + 6 + 2(0) - 1)\\ &> 1,\end{aligned}$$

the roots of Expression (1) are larger than or equal to 1.

Therefore, $\frac{\partial d_\lambda(\hat{r})}{\partial y}$ does not have roots in $y \in [0, 1)$. For each fixed θ and $\hat{r}$, by the Intermediate Value Theorem, $\frac{\partial d_\lambda(\hat{r})}{\partial y}$ does not change sign for $y \in [0, 1]$. Substituting $y = \frac{1}{2}$ confirms that $\frac{\partial d_\lambda(\hat{r})}{\partial y} \geq 0$.

Hence, the maximum of $d_\lambda(\hat{r})$ occurs when $y = 1$. Setting $y = 1$,

$$d_\lambda(\hat{r}) = \frac{2\sin\theta + \cos\theta + 2}{\sqrt{2\sin\theta + 2}}.$$

This can be elegantly optimised by substituting $\theta = \frac{\pi}{2} - 2\phi$:

$$\begin{aligned}d_\lambda(\hat{r}) &= \frac{2\cos 2\phi + \sin 2\phi + 2}{\sqrt{2 + \cos 2\phi}}\\ &= \frac{4\cos^2\phi + 2\sin\phi\cos\phi}{\sqrt{4\cos^2\phi}}\\ &= 2\cos\phi + \sin\phi\\ &= \sqrt{5}\sin(\phi + \tan^{-1} 2)\end{aligned}$$

which attains the maximum $\sqrt{5}$ when $\phi = \tan^{-1}\frac{1}{2}$ i.e. $\theta = \tan^{-1}\frac{3}{4}$. By Lemma 2.13, $s(2) \leq \sqrt{5}$, which completes the proof. □

4. Possible Approaches to n-configuration

4.1. **3-configuration.** Using graphing software, we have observed that $s(3)$ is very close to $1 + \sqrt{5}$ and we believe it is not a coincidence. However, after analysing the time complexity of constructing a computer program to check such result, we conclude that it is basically infeasible to build such program as it requires 5 layers of nested iteration i.e. $O(n^5)$. Our algorithm fixes the coordinates of a square, then for each new square, it requires its own rotation and translation on the existing configuration. Thus, each new square requires 2 layers of nested iteration. An extra iteration is required for $\hat{r}$ so there is a total of 5 layers of iterations to generate all 3-configurations. In addition, the algorithm is much more complex (refer to Section 4.2) compared to that of 2-configuration as there are more cases of the placement of the squares. Therefore, it is infeasible to find a numerical approximation of $s(n)$ for large n using this method. As a result, we investigate other approaches in an effort to generalize to n-configuration.

4.2. **"Vertex-to-Edge" and "Vertex-to-Vertex".** The major difficulty of constructing a computer program for 3-configuration is to handle the cases that the edge of a square is intersecting with a vertex or an edge of another square. Unlike 2-configuration, which only requires a single variable y to handle such cases, the newly added square in 3-configuration has multiple ways to intersect with the original configuration. It drastically increases the difficulty to design an algorithm that covers all the cases.

Thus, we try to prove that "vertex-to-edge" always results in a smaller f_λ^{min} than "vertex-to-vertex" configuration i.e. the squares only intersect at their vertices. We try to translate the square that is in "vertex-to-edge" with another square along that edge until it reaches the two end of that edge i.e. it becomes a "vertex-to-vertex".

However, when we translate a convex hull along a particular edge of another convex hull, the idea that a vertex-to-vertex connection results in a larger f_λ^{min} than a vertex-to-edge connection is not true (see the example below). For our case, which considers the boundary of a configuration rather than a convex hull, we face obstacles when trying to prove it analytically or disprove it by finding an counterexample. Although this hypothesis is too complex to be proven, it is believed that the "least optimal" configuration does not have vertex-to-edge connections according to the graphing software in both 2-configuration and 3-configuration cases. In the end, we are able to show that "vertex-to-vertex" is indeed the solution.

An approximated example will be demonstrated below. Each hollow polygon represents a convex hull.

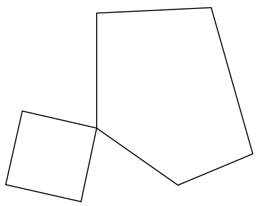


FIGURE 13. $f_\lambda^{min} = 3.070$

#	Vertices of the left polygon in Figure 13
1	$(2.701, 2.129)$
2	$(2.482, 1.153)$
3	$(3.476, 0.927)$
4	$(3.683, 1.903)$

TABLE 2

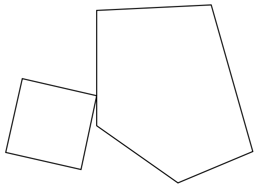


FIGURE 14. $f_\lambda^{min} = 3.095$

#	Vertices of the left polygon in Figure 14
1	$(2.701, 2.525)$
2	$(2.482, 1.549)$
3	$(3.476, 1.323)$
4	$(3.683, 2.299)$

TABLE 3

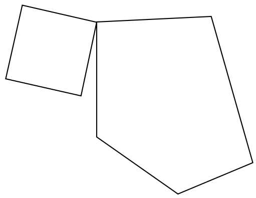


FIGURE 15. $f_\lambda^{min} = 2.919$

#	Vertices of the left polygon in Figure 15
1	$(2.701, 3.649)$
2	$(2.482, 2.673)$
3	$(3.476, 2.447)$
4	$(3.683, 3.423)$

TABLE 4

#	Vertices of the right polygon
1	$(3.683, 1.902)$
2	$(4.756, 1.146)$
3	$(3.476, 0.927)$
4	$(5.195, 3.5)$
5	$(3.683, 3.427)$

TABLE 5

5. Techniques to Tackle n-configurations

As seen in the proof of Theorem 3.3, the two parameters y and θ of an arbitrary 2-configuration already cause great algebraic complexity when proving the upper bound of $s(2)$. One can easily imagine that with the increasing number of parameters of an arbitrary n-configuration, it is neither feasible nor insightful to find $s(n)$ with a purely algebraic method.

Therefore, it is hoped that by discussing the techniques we have used to solve $s(n)$, we can offer some insights into the nature of square packing and lead into the complete proof in the next section.

5.1. Classification.

The proof of Theorem 3.3 reveals that it may be possible to only consider an α container, i.e. the case when $d_\lambda(\hat{r}) = d_\lambda(\hat{r}_\perp)$. Thus, when tackling an n-configuration, we will first categorise the configurations and containers into *classes* so that they can be solved case by case.

We now rigorously define the α container, known as class α_4.

Definition 5.1. *Let $\mathcal{C}$ be a container of a configuration λ. $\mathcal{C}$ and λ are called of class α_4 if $\mathcal{C}$ is a square, and each edge of $\mathcal{C}$ has at least one point of intersection with λ.*

Using this definition of α_4, we can rephrase Lemma 2.8 as follows.

Lemma 5.2. *For all configurations λ, there exists a container of λ of class α_4.*

Proof. By Lemma 2.8,

$$d_\lambda(\hat{r}_0) = d_\lambda(\hat{r_0}_\perp)$$

for some $\hat{r}_0 \in \mathbb{S}^1$. By the definition of d_λ, let $d_\lambda(\hat{r}_0) = d_\lambda^{ij}(\hat{r}_0)$ and $d_\lambda(\hat{r_0}_\perp) = d_\lambda^{kl}(\hat{r_0}_\perp)$ for some i, j, k, l. Then

$$d_\lambda(\hat{r}_0) = d_\lambda^{ij}(\hat{r}_0) = d_\lambda^{kl}(\hat{r_0}_\perp) = d_\lambda(\hat{r_0}_\perp),$$

so there exists a square container along the directions of $\hat{r}_0$ and $\hat{r_0}_\perp$, with point i and j on two opposite edges, and k and l on the other two opposite edges. Thus, the container and λ are of class α_4.

□

Class α_4 includes a lot of ways that the configuration and its container can intersect. Therefore, we classify α_4 into four classes $\alpha_{4.1}, \alpha_{4.2}, \alpha_{4.3}, \alpha_{4.4}$ based on the number of unit squares in the configuration which the intersection points belong to.

Definition 5.3. *For an α_4 container $\mathcal{C}$ of a configuration λ, let p, q, r, s be a point of intersection of λ and each of the edges of $\mathcal{C}$ respectively. $\mathcal{C}$ and λ are called of class $\alpha_{4.n}$, where $n = 1, 2, 3$ or 4, if p, q, r, s come from exactly n unit squares of λ.*

It should be apparent that $\alpha_{4.1}$ is impossible when the number of unit squares is greater than 1. The following simple results help prove this fact.

Lemma 5.4. *Suppose λ is an n-configuration where $n \geq 2$, and $\mathcal{C}$ and λ are of α_4. If a unit square Q in λ has a point of intersection with an edge e of $\mathcal{C}$, then Q has no point of intersection with the edge opposite to e.*

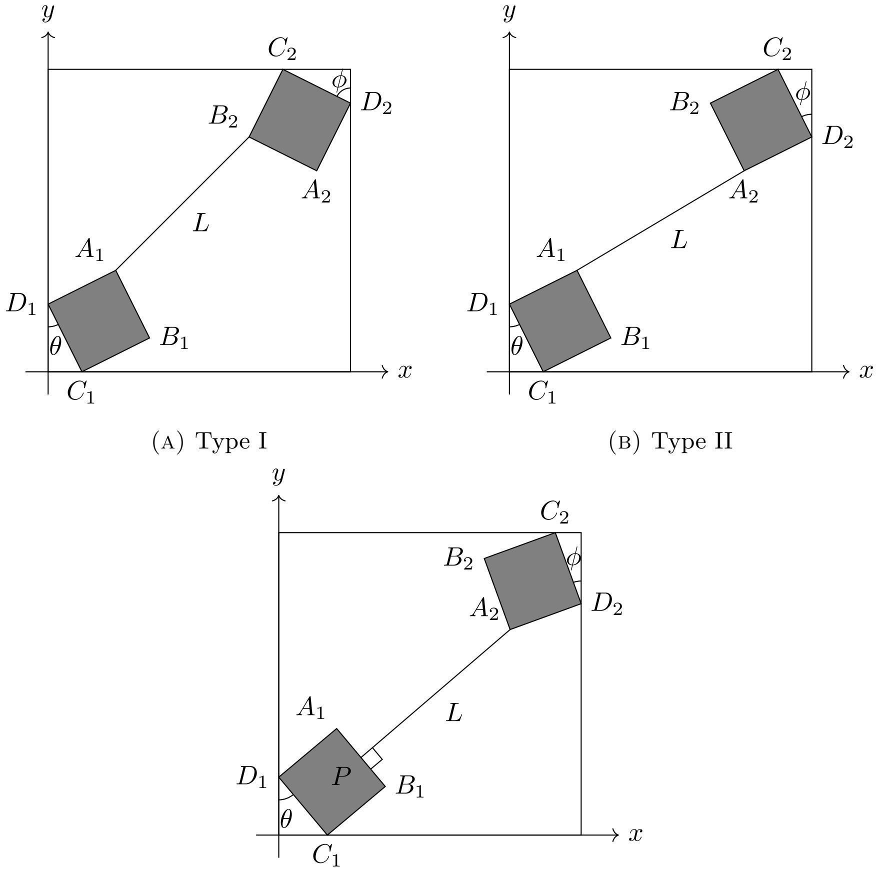
Proof. Suppose both e and the edge opposite to e each have a point of intersection with Q, denoted points x and y respectively. Then the side length of $\mathcal{C}$ is less than or equal to the length xy, which is no more than $\sqrt{2}$ since x and y are on a unit square Q. However the side length of the container of an n-configuration where $n \geq 2$ is at least 2. So Q cannot intersect with both e and its opposite edge. □

Corollary 5.5. *Suppose λ is an n-configuration where $n \geq 2$ and $\mathcal{C}$ is a container of λ. If $\mathcal{C}$ and λ are of class α_4, then they are of class $\alpha_{4.2}$, $\alpha_{4.3}$ or $\alpha_{4.4}$.*

Proof. For an $\alpha_{4.1}$ container, all four points p, q, r, s defined in Definition 5.3 come from one unit square. But by Lemma 5.4, this is impossible. □

After we divide α_4 into three smaller classes $\alpha_{4.2}$, $\alpha_{4.3}$ and $\alpha_{4.4}$, we will further subdivide $\alpha_{4.2}$ into three types.

Definition 5.6. *Suppose a configuration λ and its container $\mathcal{C}$ are of class $\alpha_{4.2}$. Let Q_1 and Q_2 be the two squares that points p, q, r, s defined in Definition 5.3 come from. For some $L \in (0, \infty)$, λ and $\mathcal{C}$ are said to be of class $\alpha_{4.2}$ type I, II or III if Q_1, Q_2 and $\mathcal{C}$ can be transformed to the respective diagram in Figure 16 for some $\theta, \phi \in [0, \pi/2]$ under translation, rotation and reflection.*



(A) Type I (B) Type II

(C) Type III

FIGURE 16. Definition of class $\alpha_{4.2}$ type I, II and III. $\theta, \phi \in [0, \pi/2]$.

Lemma 5.7. *Suppose a configuration λ and its container $\mathcal{C}$ are of class $\alpha_{4.2}$. Define Q_1 and Q_2 as in Definition 5.6. Let l be the minimum distance from a point on Q_1 to a point on Q_2. Then λ and $\mathcal{C}$ are of class $\alpha_{4.2}$ type I, II or III with $L = l$.*

Proof. By Lemma 5.4, each of Q_1 and Q_2 touches two of the adjacent sides of $\mathcal{C}$. Let p, q be the points of intersection of $\mathcal{C}$ and Q_1, and let r, s be the points of intersection of $\mathcal{C}$ and Q_2 with p, q, r, s going anticlockwise.

Let Γ be the locus of points which make a distance of l with the boundary of Q_1. Γ consists of line segments and quarter circles. By symmetry, we only consider one of those line segments as Γ_1 and one of those quarter circles as Γ_2.

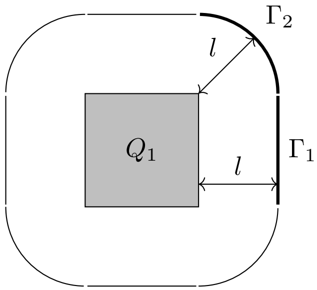
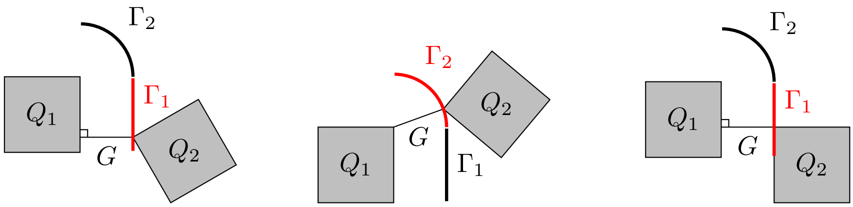


FIGURE 17. The locus Γ of points which make a distance l with the boundary of Q_1.

Q_2 has at least one point of intersection with Γ and lies on or outside Γ.

(1) If Q_2 has exactly one point of intersection with Γ, denoted as i_2,
 (a) and if i_2 is on Γ_1, then a line segment G of length l can be drawn perpendicular to Γ_1 from i_2 to the boundary of Q_1.
 (b) If i_2 is on Γ_2, then a line segment G of length l can be drawn from i_2 to the centre of Γ_2, i.e. a vertex of Q_1.

(2) If Q_2 has more than one point of intersection with Γ, then an edge of Q_2 coincides with Γ_1, and so at least one vertex v of Q_2 lies on Γ_1. Then a line segment G of length l can be drawn perpendicular to Γ_1 from v to the boundary of Q_1.



(A) Case 1(a). (B) Case 1(b). (C) Case 2.

FIGURE 18. Cases of intersection of Q_2 and Γ.

For case 1(a) and 2, λ and $\mathcal{C}$ match the definition of class $\alpha_{4.2}$ type III with $L = l$.

For case 1(b), let i_1 and i_2 be the end points of G which are on Q_1 and Q_2 respectively. p, q, r, s, i_1, i_2 are all distinct vertices of Q_1 and Q_2. Table 6 shows the corresponding type.

Vertex opposite to i_1 in Q_1	Vertex opposite to i_2 in Q_2	Type
p	r	Type II
p	s	Type I
q	r	Type I
q	s	Type II

TABLE 6

□

Figure 19 shows a summary of the classification of configurations and their containers.

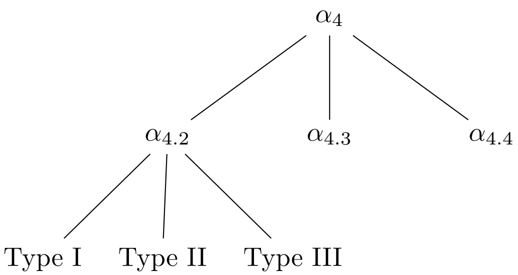


FIGURE 19. Classification of configurations and containers.

5.2. **Reducible Configurations.** After classification, the number of parameters that directly affect the container side length has reduced significantly. For instance, class $\alpha_{4.2}$ type I only has three parameters, θ, ϕ and L. However, solving the cases algebraically is still very complicated, so further techniques are required.

One intuitive thought is that for $\alpha_{4.2}$ type I, II and III, when L increases, the container side length should also increase. However, this is not always the case and the conditions for θ and ϕ when this is true are complicated to solve. If this was true, then if a container side length can be achieved by a configuration with $L < l$, we could separate Q_1 and Q_2 along the direction of l until $L = l$, which would result in a larger container side length. We call a configuration with $L = l$ "reducible" if its container side length can be achieved by a configuration with $L < l$.

Fortunately, there is an analytical method to "ignore" reducible configurations, under the condition that an upper bound $U(L)$ of the container side length of irreducible configurations increases as L. This means that $U(l)$ is also an upper bound for the container side length of reducible configurations with $L = l$. The reason behind it is that if a configuration can be reduced to one with $L = l_r$ where $l_r < l$, its container side length should be smaller than $U(l_r)$, which is no more than $U(l)$ by its increasing property.

To prove this rigorously, we first introduce a general proposition using analytical tools.

Lemma 5.8. *Let $f : [0, \infty) \to \mathbb{R}$ be a function and $g : [0, \infty) \to \mathbb{R}$ be a continuous function with the following properties:*

(a) *$g(x) \geq f(x)$ for all $x \geq 0$;*
(b) *$g(0) = f(0)$;*
(c) *for all $x_0 > 0$, if $g(x_0) \neq f(x_0)$, then there exists some $x_1 \in (0, x_0)$ such that $g(x_1) \geq g(x_0)$;*

If $U : [0, \infty) \to \mathbb{R}$ is an increasing function and is an upper bound of f, i.e. for all $x \geq 0$,

$$f(x) \leq U(x),$$

then U is also an upper bound of g, i.e. for all $x \geq 0$,

$$g(x) \leq U(x).$$

Proof. Assume $g(x_0) > U(x_0)$ for some $x_0 > 0$, so $g(x_0) > f(x_0)$. Let $X = \{x \in [0, x_0) : g(x) \geq g(x_0)\}$.

By property (c), X is non-empty. Also X is bounded since $X \subseteq [0, x_0)$. Hence X has an infimum, $\inf X$. Therefore, there exists a non-increasing sequence $s_1, s_2, s_3, \ldots \in X$ such that $\lim_{n\to\infty} s_n = \inf X$ [2].

Therefore,

$$\begin{aligned} g(\inf X) &= g(\lim_{n\to\infty} s_n) \\ &= \lim_{n\to\infty} g(s_n) && \text{using the continuity of } g \\ &\geq \lim_{n\to\infty} g(x_0) && \text{using } s_n \in X \\ &= g(x_0) \end{aligned}$$

so $\inf X \in X$. Hence X has a minimum, denoted as x_{min}.

But

$$\begin{aligned} g(x_{min}) &\geq g(x_0) \\ &> U(x_0) \\ &\geq U(x_{min}) && \text{as } U \text{ is increasing} \\ &\geq f(x_{min}) \end{aligned}$$

so $g(x_{min}) \neq f(x_{min})$. Then by property (c), there exists some $x' < x_{min}$ such that $g(x') \geq g(x_{min}) \geq g(x_0)$. $x' \in X$, which contradicts with x_{min} being the minimum of X. □

Then, we can apply this fact to our problem related to configurations.

Theorem 5.9. *Let $\sigma(l)$ be a function that returns a set of configurations for $l \in [0, \infty)$. Suppose $\sigma(l)$ is disjoint from $\sigma(l')$ for any $l' \neq l$. Let $f : \Lambda \to \mathbb{R}$. For any $l \in [0, \infty)$, $\sigma_r(l) \subset \sigma(l)$ is a set of "reducible configurations", i.e. for any $\lambda \in \sigma_r(l)$, $f(\lambda) = f(\lambda_r)$ where $\lambda_r \in \sigma(l_r)$ and $l_r \in (0, l)$. Also assume two functions*

$$m(l) = \max\{f(\lambda) : \lambda \in \sigma(l)\}, \quad m_r(l) = \max\{f(\lambda) : \lambda \in \sigma(l) \setminus \sigma_r(l)\}$$

are well-defined.

Let $U(l)$ be an upper bound of $m_r(l)$. If $m(l)$ is continuous and $U(l)$ is increasing, then $U(l)$ is also an upper bound of $m(l)$.

Proof. We will confirm the three properties of Lemma 5.8:

(a) Since $\sigma(l) \setminus \sigma_r(l) \subseteq \sigma(l)$, $m(l) \geq m_r(l)$.
(b) From the definition of σ_r, $\sigma_r(0) = \emptyset$. Therefore $\sigma(0)\backslash\sigma_r(0) = \sigma(0)\backslash\emptyset = \sigma(0)$ so $m(0) = m_r(0)$.
(c) For all $l_0 > 0$, using the definition of m, let $\lambda \in \sigma(l_0)$ where $f(\lambda) = m(l_0)$.
If $m(l_0) \neq m_r(l_0)$, then $m(l_0) > m_r(l_0)$. Since $f(\lambda) > m_r(l_0)$, by definition of m_r, $\lambda \notin \sigma(l_0) \setminus \sigma_r(l_0)$ so $\lambda \in \sigma_r(l_0)$.
By definition of σ_r, let $f(\lambda) = f(\lambda_r)$ where $\lambda_r \in \sigma(l_r)$ and $l_r \in (0, l_0)$. Therefore by definition of m, $m(l_r) \geq f(\lambda_r) = f(\lambda) = m(l_0)$.

Hence, $U(l)$ is an upper bound of $m(l)$. □

Corollary 5.10. *Define m and m_r as in Theorem 5.9. If m_r is an increasing function, then for all $l \in [0, \infty)$,*

$$m(l) = m_r(l).$$

Proof. $m_r(l)$ is an upper bound of itself, so let $U(l) = m_r(l)$. By Theorem 5.9, $m(l) \leq U(l) = m_r(l)$ for all $l \in [0, \infty)$. But since also $m(l) \geq m_r(l)$, we have $m(l) = m_r(l)$. □

The above two results are powerful tools to use the concept of "reducible configurations". In our later proof, we will first assume a configuration is not reducible. Next, we will show that the container side length of a configuration that satisfies this assumption has an upper bound which increases with L. Theorem 5.9 or Corollary 5.10 then justifies the assumption since reducible configurations can be ignored.

5.3. **Local Minimum.** Now that we have developed tools to ignore reducible configurations, such configurations should be identified to further lower the number of parameters.

For class $\alpha_{4.2}$ type I, II or III, suppose θ and ϕ are not 0 or $\pi/2$. Fixing the container side length, consider L as a function of θ and ϕ. If L is not at a local minimum when changing θ or ϕ, then a slight increase or decrease in θ and/or ϕ results in a smaller L.

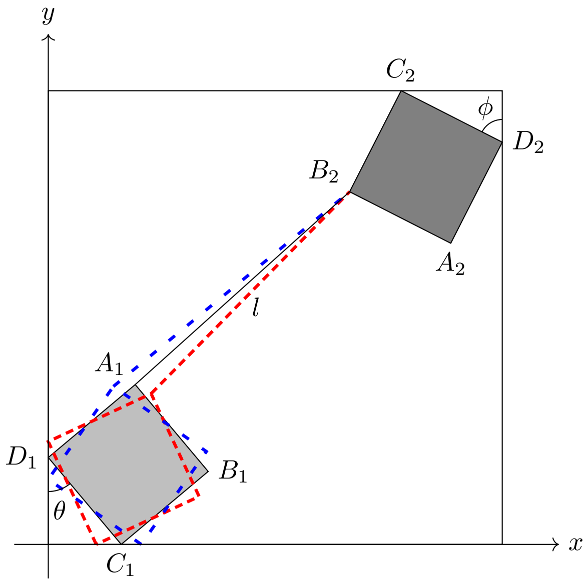


FIGURE 20. If L is not at a local minimum when changing θ, then the configuration is reducible. A slight decrease in θ, shown as red densely dashed lines, has a smaller length of A_1B_2.

Hence, we are interested in finding the conditions when L is at a local minimum.

Lemma 5.11. *Suppose $v(t)$ and u are two vectors in $\mathbb{R}^2$ where $v \neq u$ and v is parametrised by $t \in \mathbb{R}$. Then $\|v-u\|$ attains a local minimum only if*

$$\frac{dv}{dt} \cdot (v-u) = 0.$$

Proof. At a local minimum,

$$\begin{aligned} 0 &= \frac{d}{dt}\|v-u\| \\ &= \frac{(v-u)\cdot \frac{dv}{dt}}{\|v-u\|} \end{aligned}$$

so $\frac{dv}{dt} \cdot (v-u) = 0$. □

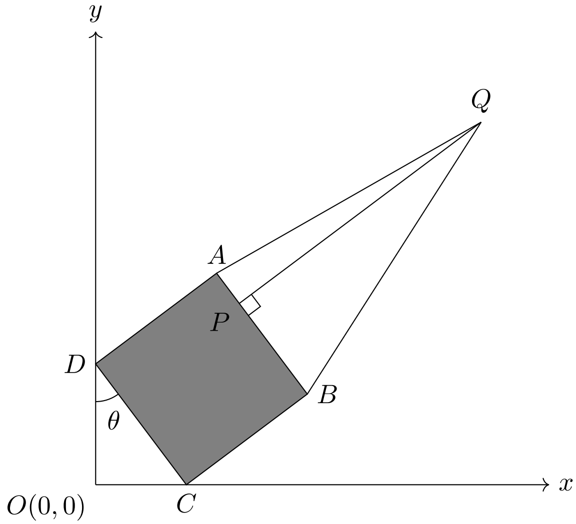


FIGURE 21. A unit square $ABCD$ which makes a variable angle θ with the container, a fixed point Q, and a point P on AB such that $PQ \perp AB$.

Lemma 5.12. *In Figure 21, for a fixed point Q,*

(a) *length AQ attains a local minimum at $\theta = \theta_0$ only if AQ is along* $(\cos\theta_0 - \sin\theta_0, \sin\theta_0)$*;*
(b) *length BQ attains a local minimum at $\theta = \theta_0$ only if BQ is along* $(\cos\theta_0, \sin\theta_0 - \cos\theta_0)$*; and*
(c) *length PQ attains a local minimum at $\theta = \theta_0$ only if $AP = \sin^2\theta_0$ and $PQ \leq 3\sin\theta_0\cos\theta_0 - 1$.*

Proof. (a) The position vector of $A = (\cos\theta, \sin\theta + \cos\theta)$ has derivative

$$\frac{d}{d\theta}(\cos\theta, \sin\theta + \cos\theta) = (-\sin\theta, \cos\theta - \sin\theta)$$

so by Lemma 5.11, AQ is perpendicular to the derivative, i.e. along $(\cos\theta_0 - \sin\theta_0, \sin\theta_0)$.

(b) The position vector of $B = (\sin\theta + \cos\theta, \sin\theta)$ has derivative

$$\frac{d}{d\theta}(\sin\theta + \cos\theta, \sin\theta) = (\cos\theta - \sin\theta, \cos\theta)$$

so by Lemma 5.11, BQ is perpendicular to the derivative, i.e. along $(\cos\theta_0, \sin\theta_0 - \cos\theta_0)$.

(c) Let y be the length of AP and l be the length of PQ. Furthermore let y_0 and l_0 be the values of y and l respectively when $\theta = \theta_0$. The position vector of Q is

$$OQ = (y_0\sin\theta_0 + (1 + l_0)\cos\theta_0, (1 + l_0)\sin\theta_0 + (1 - y_0)\cos\theta_0),$$

and

$$\begin{aligned} l &= (\cos\theta, \sin\theta) \cdot (OQ - (\cos\theta, \sin\theta + \cos\theta)) \\ &= OQ \cdot (\cos\theta, \sin\theta) - 1 - \sin\theta\cos\theta. \end{aligned}$$

We now compute its derivatives,

$$\frac{dl}{d\theta} = OQ \cdot (-\sin\theta, \cos\theta) + \sin^2\theta - \cos^2\theta$$

and

$$\frac{d^2l}{d\theta^2} = -OQ \cdot (\cos\theta, \sin\theta) + 4\sin\theta\cos\theta.$$

Since $l = l_0$ is a local minimum,

$$\begin{aligned} 0 &= \left.\frac{dl}{d\theta}\right|_{\theta=\theta_0} \\ &= -y_0\sin^2\theta_0 - (1+l_0)\sin\theta_0\cos\theta_0 + (1+l_0)\sin\theta_0\cos\theta_0 \\ &\quad + (1-y_0)\cos^2\theta_0 + \sin^2\theta_0 - \cos^2\theta_0 \\ y_0 &= \sin^2\theta_0. \end{aligned}$$

Besides, $\frac{d^2l}{d\theta^2} < 0$ would imply $l = l_0$ is a local maximum instead, so a local minimum requires

$$\begin{aligned} \left.\frac{d^2l}{d\theta^2}\right|_{\theta=\theta_0} &\geq 0 \\ -(l_0 + 1 + \sin\theta_0\cos\theta_0) + 4\sin\theta_0\cos\theta_0 &\geq 0 \\ l_0 &\leq 3\sin\theta_0\cos\theta_0 - 1. \end{aligned}$$

□

5.4. **Path Rearrangement.** Our final technique involves simplifying a sequence of unit squares and rearranging them. Since in our definition of a configuration, all unit squares are connected to each other, we can obtain a sequence of connected unit squares from any starting square to any destination square, which we call a path.

In most of the cases, the details of the path do not affect the side length of the container, so it can be simplified. Besides, squares in a path can be rearranged in a way to transform into a different class without affecting the container side length. The implementation of this technique will be demonstrated when tackling class $\alpha_{4.3}$ and $\alpha_{4.4}$.

6. Complete Proof

We now apply the techniques mentioned in the previous section to find the maximum container side length of configurations. Different classes will be handled in each subsection.

6.1. $\alpha_{4.2}$ **Type I.** For $\alpha_{4.2}$ type I, we mainly use the fact that L is at a local minimum when changing θ and ϕ. However, we need to first handle the case when θ or ϕ is 0 or $\pi/2$. In this case, since $\theta \in [0, \pi/2]$ or $\phi \in [0, \pi/2]$ is at an end of its range, the notion of local minima does not apply. Luckily, this is easy to tackle using algebraic methods.

Lemma 6.1. *The side length of a container of class $\alpha_{4.2}$ type I or III with $L = l$ and $\phi = \pi/2$ is at most*

$$\frac{l}{\sqrt{2}} + \frac{\sqrt{5}}{2} + 1.$$

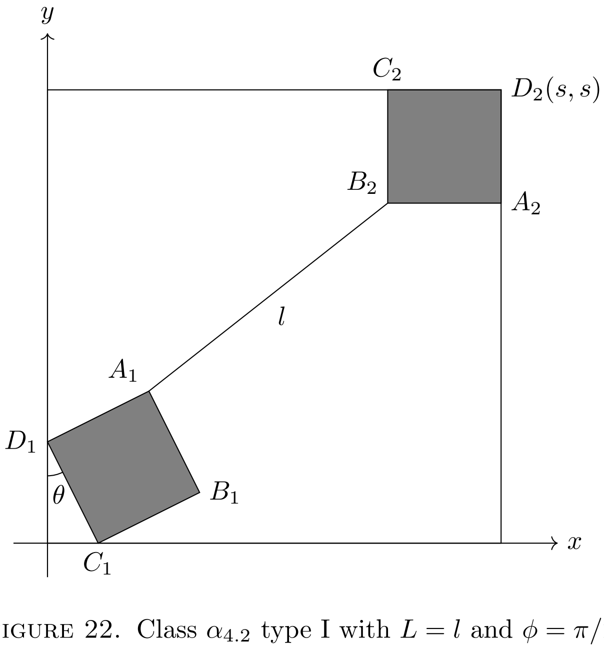


FIGURE 22. Class $\alpha_{4.2}$ type I with $L = l$ and $\phi = \pi/2$.

Proof. Let s be the side length of the container. The top right vertex of the container has coordinates (s, s).

For a configuration of class $\alpha_{4.2}$ type I, A_1 has coordinates $(\cos\theta, \sin\theta + \cos\theta)$. For a configuration of class $\alpha_{4.2}$ type III, P has coordinates $(\cos\theta + t\sin\theta, \sin\theta + (1-t)\cos\theta)$ where t is the distance A_1P. So by considering the length of l, for some $t \in [0,1]$,

$$\begin{aligned}
l^2 &= \big((\cos\theta + t\sin\theta) - (s-1)\big)^2 + \big((\sin\theta + (1-t)\cos\theta) - (s-1)\big)^2 \\
&= (\cos\theta + t\sin\theta)^2 + (\sin\theta + (1-t)\cos\theta)^2 \\
&\quad - 2((1+t)\sin\theta + (2-t)\cos\theta)(s-1) + 2(s-1)^2 \\
&= \Big(\frac{(1+t)\sin\theta + (2-t)\cos\theta}{\sqrt{2}} - \sqrt{2}(s-1)\Big)^2 + \Big(\frac{(t-1)\sin\theta - t\cos\theta}{\sqrt{2}}\Big)^2 \\
&\geq \Big(\frac{(1+t)\sin\theta + (2-t)\cos\theta}{\sqrt{2}} - \sqrt{2}(s-1)\Big)^2.
\end{aligned}$$

Taking the square roots,

$$\begin{aligned}
-l &\leq \frac{(1+t)\sin\theta + (2-t)\cos\theta}{\sqrt{2}} - \sqrt{2}(s-1) \\
s &\leq \frac{l}{\sqrt{2}} + \frac{(1+t)\sin\theta + (2-t)\cos\theta}{2} + 1 \\
&\leq \frac{l}{\sqrt{2}} + \frac{\sqrt{(1+t)^2 + (2-t)^2}}{2} + 1 \\
&= \frac{l}{\sqrt{2}} + \frac{\sqrt{5 - 2t(1-t)}}{2} + 1 \\
&\leq \frac{l}{\sqrt{2}} + \frac{\sqrt{5}}{2} + 1.
\end{aligned}$$

□

Now we are ready to find the maximum side length of a container of class $\alpha_{4.2}$ type I. Two conditions will be used, namely L being at a local minimum, and that the container is a square. After applying these conditions, an equation will be obtained, which can be solved algebraically.

Lemma 6.2. $\alpha_{4.2}$ *Type I with* $L = l$ *has a maximum container side length of*

$$\frac{l}{\sqrt{2}} + \sqrt{5}.$$

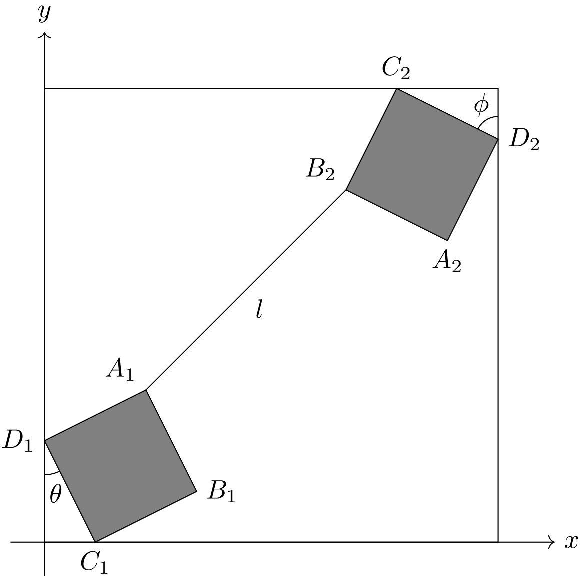


FIGURE 23. Class $\alpha_{4.2}$ type I.

Proof. Consider class $\alpha_{4.2}$ type I with maximum container side length. By Lemma 6.1, when $\phi = \pi/2$, or, by rotational symmetry, $\theta = 0$, the container side length is at most

$$\frac{l}{\sqrt{2}} + \frac{\sqrt{5}}{2} + 1.$$

Also $\theta = \pi/2$ or $\phi = 0$ is impossible since if so, A_1B_2 is not the shortest distance from Q_1 to Q_2. From now on, we will focus on the case that θ and ϕ are not 0 or $\pi/2$.

We first assume the condition that the length L is at a local minimum when changing θ and ϕ. We will later justify this assumption using Corollary 5.10.

Fixing ϕ and therefore point B_2, since the length A_1B_2 attains a local minimum at this θ, by Lemma 5.12, A_1B_2 is along $(\cos\theta - \sin\theta, \sin\theta)$. Similarly by fixing θ and considering ϕ, A_1B_2 is along $(\cos\phi, \sin\phi - \cos\phi)$. Therefore

$$\begin{aligned}
\vec{0} &= (\cos\theta - \sin\theta, \sin\theta) \times (\cos\phi, \sin\phi - \cos\phi) \\
\sin\theta \sin\phi &= \cos\theta(\sin\phi - \cos\phi) \\
\tan\phi &= \frac{1}{1 - \tan\theta} \\
\cos\phi &= \frac{1 - \tan\theta}{\sqrt{(1 - \tan\theta)^2 + 1}}.
\end{aligned}$$

Let $\vec{v}_{min} = (\cos\theta - \sin\theta, \sin\theta)$. The top right vertex of the container has the coordinates

$$\left(\cos\theta + \sin\phi + \cos\phi + \frac{l}{\|\vec{v}_{min}\|}(\cos\theta - \sin\theta), \sin\theta + \cos\theta + \sin\phi + \frac{l}{\|\vec{v}_{min}\|}\sin\theta\right).$$

Since the container is a square, the x and y-coordinates should equal. Hence

$$\frac{l}{\|\vec{v}_{min}\|}(\cos\theta - 2\sin\theta) = \sin\theta - \cos\phi. \tag{2}$$

$\frac{l}{\|\vec{v}_{min}\|} > 0$, so if $(\cos\theta - 2\sin\theta)$ and $(\sin\theta - \cos\phi)$ are not both zero, they have the same sign. However, we will now show that this is not possible.

Consider when $\sin\theta - \cos\phi > 0$,

$$\begin{aligned} \sin\theta &> \frac{1-\tan\theta}{\sqrt{(1-\tan\theta)^2+1}} \\ \sin\theta\sqrt{(1-\tan\theta)^2+1} &> 1-\tan\theta. \end{aligned}$$

For $\pi/4 < \theta < \pi/2$, this is always true since the left hand side is positive while the right hand side is negative. For $0 \le \theta \le \pi/4$,

$$\begin{aligned} \sin^2\theta((1-\tan x)^2+1) &> (1-\tan\theta)^2 \\ \sin^2\theta &> \cos^2\theta(1-\tan\theta)^2 \\ \tan^2\theta &> (1-\tan\theta)^2 \\ \tan\theta &> 1-\tan\theta \\ \theta &> \tan^{-1}\frac{1}{2}. \end{aligned}$$

Therefore, $(\sin\theta - \cos\phi)$ is positive if and only if $\theta > \tan^{-1}\frac{1}{2}$. In the meantime, $(\cos\theta - 2\sin\theta)$ is negative if and only if $\theta > \tan^{-1}\frac{1}{2}$.

Hence, Equation (2) is only true when $\cos\theta - 2\sin\theta = 0$. Solving yields $\theta = \tan^{-1}\frac{1}{2}$ and $\phi = \tan^{-1} 2$. Substituting the angles, $\|\vec{v}_{min}\| = \sqrt{2/5}$ and the side length of the container is

$$\frac{l}{\sqrt{2}} + \sqrt{5}.$$

Note that this is an increasing function of l.

Now we will justify the assumption that L is at a local minimum when changing θ and ϕ. If L is not at a local minimum, then by slightly increasing or decreasing θ and/or ϕ, L decreases to smaller than l while keeping the side length of the container unchanged. Hence by Corollary 5.10, this is a "reducible" configuration and does not affect the maximum side length of the container. □

6.2. $\alpha_{4.2}$ **Type II.** For class $\alpha_{4.2}$ type II, it can be reduced to class $\alpha_{4.2}$ type I. Again, using the conditions of local minima and being a square, a relationship between θ, ϕ, l and the container side length can be found. Finally, it can be shown that reflecting one of the unit squares results in a configuration of class $\alpha_{4.2}$ type I.

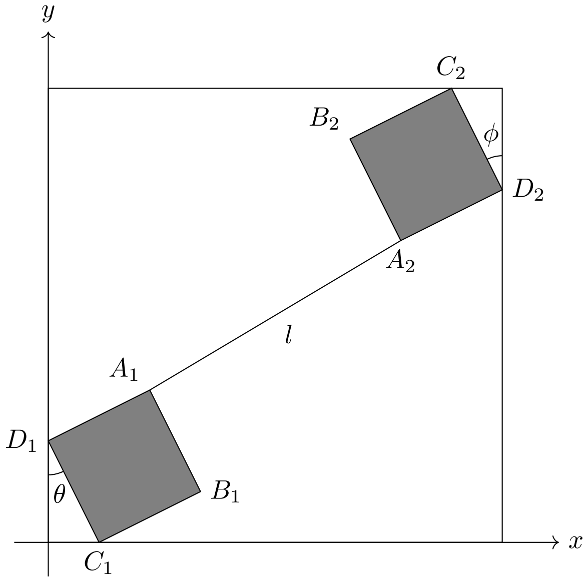


FIGURE 24. Class $\alpha_{4.2}$ type II.

Lemma 6.3. *Any $\alpha_{4.2}$ type II with $L = l$ has a smaller side length of container than that of class $\alpha_{4.2}$ type I also with $L = l$.*

Proof. Consider an $\alpha_{4.2}$ type II with maximum container side length. If $\theta = 0$ or $\phi = 0$, by Lemma 6.1, the container side length is at most

$$\frac{l}{\sqrt{2}} + \frac{\sqrt{5}}{2} + 1,$$

which is smaller than the maximum side length of a container of $\alpha_{4.2}$ type I. Also $\theta = \pi/2$ or $\phi = \pi/2$ is impossible since if so, A_1A_2 is not the shortest distance between Q_1 and Q_2. Therefore from now on, we assume θ and ϕ are not 0 or $\pi/2$.

By Theorem 5.9, we first assume this configuration is not reducible. Then we can assume that L is at a local minimum when changing θ and ϕ. This condition, by Lemma 5.12, requires A_1A_2 to be along both $(\cos\theta - \sin\theta, \sin\theta)$ and $(\cos\phi - \sin\phi, \sin\phi)$. Thus

$$\begin{aligned} \vec{0} &= (\cos\theta - \sin\theta, \sin\theta) \times (\cos\phi - \sin\phi, \sin\phi) \\ 0 &= \cos\theta\sin\phi - \sin\theta\sin\phi - \sin\theta\cos\phi + \sin\theta\sin\phi \\ &= \sin(\phi - \theta) \end{aligned}$$

so $\theta = \phi$.

Let $\vec{v}_{min} = (\cos\theta - \sin\theta, \sin\theta)$ and s be the side length of the container. The top right vertex of the container has the coordinates

$$(s,s) = \Big(\cos\theta + \cos\phi + \frac{l}{\|\vec{v}_{min}\|}(\cos\theta - \sin\theta),\\ \sin\theta + \cos\theta + \sin\phi + \cos\phi + \frac{l}{\|\vec{v}_{min}\|}\sin\theta\Big).$$

Since the container is a square, the x and y-coordinates should equal. Hence

$$\frac{l}{\|\vec{v}_{min}\|}(\cos\theta - 2\sin\theta) = 2\sin\theta.$$

Since $\frac{l}{\|\vec{v}_{min}\|} > 0$, $\cos\theta - 2\sin\theta > 0$ and

$$\frac{l}{\|\vec{v}_{min}\|} = \frac{2\sin\theta}{\cos\theta - 2\sin\theta}.$$

Substituting, the side length of the container s is

$$\begin{aligned} s &= 2\cos\theta + \frac{2\sin\theta}{\cos\theta - 2\sin\theta}(\cos\theta - \sin\theta) \\ &= \frac{2\cos^2\theta - 2\sin\theta\cos\theta - 2\sin^2\theta}{\cos\theta - 2\sin\theta}. \end{aligned}$$

Now we will show that a container of $\alpha_{4.2}$ type II is always smaller than an $\alpha_{4.2}$ type I with $L = l$. Notice that one of the unit square can be reflected about the line $y = x$ without changing the size of the container.

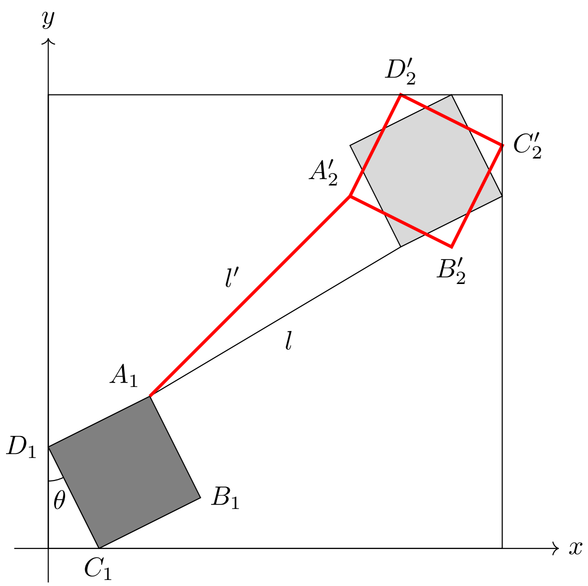


FIGURE 25. $A_2'B_2'C_2'D_2'$ is the image of the reflection of square $A_2B_2C_2D_2$ along the line $y = x$.

The length A_1A_2' is denoted l'. Note that the configuration that includes $\{A_1, B_1, C_1, D_1, A_2', B_2', C_2', D_2'\}$ with $L = l'$ is of class $\alpha_{4.2}$ type I. l' can be expressed as

$$\begin{aligned} l' &= \|(s,s) - (\cos\theta, \sin\theta + \cos\theta) - (\sin\phi + \cos\phi, \cos\phi)\| \\ &= \|(s - 2\cos\theta - \sin\theta, s - 2\cos\theta - \sin\theta)\| \\ &= \sqrt{2}(s - 2\cos\theta - \sin\theta) \\ &= \sqrt{2}\Big(\frac{2\cos^2\theta - 2\sin\theta\cos\theta - 2\sin^2\theta}{\cos\theta - 2\sin\theta} - 2\cos\theta - \sin\theta\Big) \\ &= \frac{\sqrt{2}\sin\theta\cos\theta}{\cos\theta - 2\sin\theta} \\ &= \frac{\cos\theta}{\sqrt{2}\|\vec{v}_{min}\|} l, \end{aligned}$$

but since

$$\begin{aligned} (\sqrt{2}\|\vec{v}_{min}\|)^2 - (\cos\theta)^2 &= 2\big((\cos\theta - \sin\theta)^2 + \sin^2\theta\big) - \cos^2\theta \\ &= (\cos\theta - 2\sin\theta)^2 \\ &> 0, \end{aligned}$$

we have $\cos\theta < \sqrt{2}\|\vec{v}_{min}\|$, so

$$l' < l.$$

By Lemma 6.2, the side length of this container is at most $(\frac{l'}{\sqrt{2}} + \sqrt{5})$, so

$$s \le \frac{l'}{\sqrt{2}} + \sqrt{5} < \frac{l}{\sqrt{2}} + \sqrt{5}$$

and hence such $\alpha_{4.2}$ type II has a container side length smaller than the $\alpha_{4.2}$ type I with $L = l$ that has maximum container side length. □

6.3. $\alpha_{4.2}$ **Type III.** Next, a similar technique can be applied to class $\alpha_{4.2}$ type III. However, we will soon prove that this class cannot simultaneously satisfy the conditions of local minima and the container being a square.

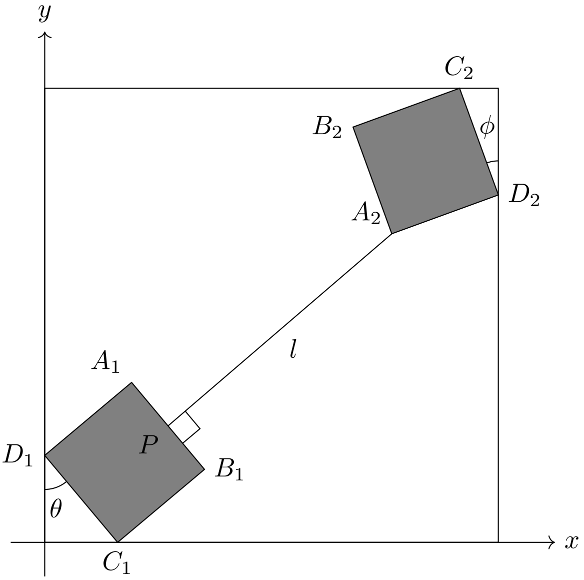


FIGURE 26. Class $\alpha_{4.2}$ type III.

Lemma 6.4. *Any $\alpha_{4.2}$ type III with $L = l$ has a smaller side length of container than that of $\alpha_{4.2}$ type I also with $L = l$.*

Proof. Consider an $\alpha_{4.2}$ type III with maximum container side length. Since $PA_2 \perp A_1B_1$, PA_2 is along $(\cos\theta, \sin\theta)$. By Theorem 5.9, assume this configuration is not reducible. Again by Lemma 6.1, if $\phi = \pi/2$, the container is smaller than that of a container of class $\alpha_{4.2}$ type I. So from now on, we will consider when θ and ϕ are not 0 and $\pi/2$. Then we can assume that L is at a local minimum when changing θ and ϕ. For being a local minimum with respect to ϕ, Lemma 5.12 requires A_2P to be also along $(\cos\phi - \sin\phi, \sin\phi)$. Thus

$$\begin{aligned}
\vec{0} &= (\cos\theta, \sin\theta) \times (\cos\phi - \sin\phi, \sin\phi) \\
\cos\theta \sin\phi &= \sin\theta(\cos\phi - \sin\phi) \\
\tan\phi &= \frac{1}{\frac{1}{\tan\theta} + 1} \\
\sin\phi &= \frac{1}{\sqrt{1 + \left(1 + \frac{1}{\tan\theta}\right)^2}}.
\end{aligned}$$

Besides, the condition that L is at a local minimum with respect to θ requires, by Lemma 5.12, that $A_1P = \sin^2\theta$ and $l \leq 3\sin\theta\cos\theta - 1$.

Combining the results, the top right vertex of the container has coordinates

$$\begin{aligned}&\big(\sin^2\theta\sin\theta + (1+l)\cos\theta + \cos\phi,\\ &(1+l)\sin\theta + (1-\sin^2\theta)\cos\theta + \sin\phi + \cos\phi\big)\\ =&\big(\sin^3\theta + (1+l)\cos\theta + \cos\phi, (1+l)\sin\theta + \cos^3\theta + \sin\phi + \cos\phi\big).\end{aligned}$$

The x and y coordinates should equal:

$$\begin{aligned}\sin^3\theta + (1+l)\cos\theta + \cos\phi &= (1+l)\sin\theta + \cos^3\theta + \sin\phi + \cos\phi\\ \sin\phi &= (\sin\theta - \cos\theta)(\sin\theta\cos\theta - l)\\ 1 &= (\sin\theta - \cos\theta)(\sin\theta\cos\theta - l)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}.\end{aligned}$$

We will show that this equation has no solution for θ using the condition

$$0 \le l \le 3\sin\theta_0\cos\theta_0 - 1.$$

For $\sin\theta - \cos\theta < 0$, i.e. $0 \le \theta < \pi/4$,

$$\begin{aligned}&(\sin\theta - \cos\theta)(\sin\theta\cos\theta - l)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}\\ \le&(\sin\theta - \cos\theta)\big(\sin\theta\cos\theta - (3\sin\theta\cos\theta - 1)\big)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}\\ =&(\sin\theta - \cos\theta)\big(1 - \sin 2\theta\big)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}\\ \le& 0 < 1.\end{aligned}$$

For $\sin\theta - \cos\theta \ge 0$, i.e. $\pi/4 \le \theta < \pi/2$,

$$\begin{aligned}&(\sin\theta - \cos\theta)(\sin\theta\cos\theta - l)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}\\ \le&(\sin\theta - \cos\theta)(\sin\theta\cos\theta)\sqrt{1 + \big(1 + \frac{1}{\tan\theta}\big)^2}\\ \le&\sqrt{5}(\sin\theta - \cos\theta)\sin\theta\cos\theta\\ \le&\sqrt{5}\sin^2\theta\cos\theta\\ \le&\frac{2\sqrt{15}}{9} < 1.\end{aligned}$$

Therefore, no $\alpha_{4.2}$ type III satisfies the local minima condition. □

6.4. $\alpha_{4.3}$ **and** $\alpha_{4.4}$. Our techniques for solving class $\alpha_{4.3}$ and $\alpha_{4.4}$ are different from $\alpha_{4.2}$. We will mostly use the path rearrangement method.

Before we tackle $\alpha_{4.3}$ and $\alpha_{4.4}$, a simple case, denoted $\alpha^{\star}_{4.3}$, is first considered to ensure a smoother presentation later on.

Lemma 6.5. *For some* $l_1, l_2 \ge 0$ *and* $\theta \in [0, \pi/2]$, *if* $l_1 + l_2 \le (n-1)\sqrt{2}$ *where* $n \ge 2$, *then the container side length of class* $\alpha^{\star}_{4.3}$, *defined in Figure 27, is less than*

$$n + \sqrt{5} - 2.$$

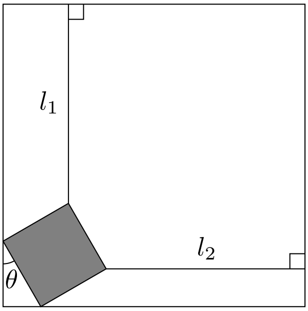


FIGURE 27. A special case of $\alpha_{4.3}$, denoted as $\alpha^\star_{4.3}$.

Proof. To form a square,

$$\begin{aligned} l_1 + \sin\theta + \cos\theta &= l_2 + \sin\theta + \cos\theta \\ l_1 &= l_2. \end{aligned}$$

Notice,

$$\begin{aligned} n - 1 &\geq \frac{l_1}{\sqrt{2}} + \frac{l_2}{\sqrt{2}} \\ &= \sqrt{2} l_1. \end{aligned}$$

Thus the side length of the square

$$\begin{aligned} &= \sin\theta + \cos\theta + l_1 \\ &\leq \sin\theta + \cos\theta + \frac{n-1}{\sqrt{2}} \\ &\leq \sqrt{2} + \frac{n-1}{\sqrt{2}} \\ &= \frac{n+1}{\sqrt{2}} \\ &< n + \sqrt{5} - 2 \quad \text{for } n \geq 2. \end{aligned}$$

□

Class $\alpha_{4.3}$ can now be dealt with. We will first construct paths between particular squares of interest, and then deduce cases based on the way which the paths split. For each case, we perform a series of steps to transform it into either class $\alpha^\star_{4.3}$ or class $\alpha_{4.2}$ while carefully considering inequalities related to triangles and the total number of unit squares.

Lemma 6.6. *The side length of the square container of any n-configuration of $\alpha_{4.3}$ is always smaller than that of $\alpha_{4.2}$.*

Proof. Suppose p, q, r, s are points of intersection of an n-configuration and each of the four edges of its square container of class $\alpha_{4.3}$. Let e_p, e_q, e_r, e_s be the container edges which p, q, r, s are on respectively. Without loss of generality, let Q_0 be the unit square that contains p and q, Q_1 be the unit square that contains r, and Q_2

be the unit square that contains s. p and q, by Lemma 5.4, are on the adjacent sides of the container and are adjacent vertices of Q_0.

Since all squares in a configuration are connected, let K_1 be the finite sequence of all distinct squares on a path starting from Q_0 and ending at Q_1, including Q_0 and Q_1. Similarly let K_2 be the finite sequence of all distinct squares on a path starting from Q_0 and ending at Q_2, including Q_0 and Q_2.

Let Q_c be the last square in K_2 which is also in K_1. Q_c exists because at least one square, Q_0, is in both K_1 and K_2.

There are three cases for Q_c:

(a) $Q_c = Q_0$;
(b) $Q_c = Q_1$; or
(c) otherwise, i.e. other squares in K_1 excluding Q_0 and Q_1.

For case (a) (Figure 28), we will show that the configuration can be transformed to be of class $\alpha^\star_{4.3}$ defined in Lemma 6.5. Let i_{01} be the intersection point of Q_0 and its next square in K_1 and i_{02} be the intersection point of Q_0 and its next square in K_2. Draw a line segment, denoted G_r with length l_r, from i_{01} to r. Also draw another line segment, denoted G_s with length l_s, from i_{02} to s.

Notice that G_r represents at most $\lceil l_r/\sqrt{2}\rceil$ unit squares and G_s represents at most $\lceil l_s/\sqrt{2}\rceil$ unit squares. Hence, if we count the total number of unit squares, we have

$$\left\lceil \frac{l_r}{\sqrt{2}} \right\rceil + \left\lceil \frac{l_s}{\sqrt{2}} \right\rceil + 1 \leq n.$$

As $l_r/\sqrt{2} \leq \lceil l_r/\sqrt{2}\rceil$ and $l_s/\sqrt{2} \leq \lceil l_s/\sqrt{2}\rceil$, we can loosen this condition to

$$l_r + l_s \leq (n-1)\sqrt{2}$$

which still holds.

Now let $i_{01\perp}$ be a point on Q_0 with shortest distance to e_r, and let $i_{02\perp}$ be a point on Q_0 with shortest distance to e_s. Draw a line segment, denoted $G_{r\perp}$ with length $l_{r\perp}$, perpendicular to e_r, from $i_{01\perp}$ to e_r. Also draw a line segment, denoted $G_{s\perp}$ with length $l_{s\perp}$, perpendicular to e_s, from $i_{02\perp}$ to e_s. By definition, $l_{r\perp} \leq l_r$ and $l_{s\perp} \leq l_s$, so

$$l_{r\perp} + l_{s\perp} \leq (n-1)\sqrt{2}.$$

Note that the side length of the container stays unchanged, while Q_0, $l_{r\perp}$ and $l_{s\perp}$ match the diagram of class $\alpha^\star_{4.3}$. Hence by Lemma 6.5, the container side length is less than

$$n + \sqrt{5} - 2,$$

which is the maximum container side length of class $\alpha_{4.2}$.

For cases (b) (Figure 29) and (c) (Figure 30), we will show that the container side length can be achieved by a configuration of class $\alpha_{4.2}$ with $L \leq (n-2)\sqrt{2}$. Then by Lemma 6.2, the container side length is no more than

$$\frac{(n-2)\sqrt{2}}{\sqrt{2}} + \sqrt{5} = n + \sqrt{5} - 2,$$

thus completing the proof.

For case (b), refer to Figure 29. Let i_{0c} be the intersection point of Q_0 and its next square in K_1, i_{c0} be the intersection point of Q_c and its previous square in K_1, and i_{c2} be the intersection point of Q_c and its next square in K_2. Draw a line segment, denoted G_{pq} with length l_{pq}, from i_{0c} to i_{c0}. Besides, draw another line segment, denoted G_s with length l_s, from i_{c2} to s.

Note that G_{pq} represents at most $\lceil l_{pq}/\sqrt{2}\rceil$ squares and G_s represents at most $\lceil l_s/\sqrt{2}\rceil$ squares. Thus counting the total number of squares, we have

$$\left\lceil \frac{l_{pq}}{\sqrt{2}} \right\rceil + \left\lceil \frac{l_s}{\sqrt{2}} \right\rceil + 2 \leq n.$$

We loosen this condition to

$$l_{pq} + l_s \leq (n-2)\sqrt{2}$$

which still holds.

We perform a series of transformations to the diagram as follows:

(1) Let $i_{c2\perp}$ be a point on Q_c with shortest distance to e_s. Draw a line segment, denoted $G_{s\perp}$ with length $l_{s\perp}$, perpendicular to e_s, from $i_{c2\perp}$ to e_s.
  Note that $l_{s\perp} \leq l_s$, so $l_{pq} + l_{s\perp} \leq (n-2)\sqrt{2}$ holds.
(2) Swap the positions of Q_c and $G_{s\perp}$, i.e. translate $G_{s\perp}$ such that $G_{s\perp}$ attaches at the end of G_{pq}, and translate Q_c and i_{c0} together such that the ending point of $G_{s\perp}$ is point i_{c0}.
  Q_c now touches both e_r and e_s.
(3) Draw a line segment, denoted G_+ with length l_+, from i_{0c} to i_{c0}.
  By the triangle inequality,

$$l_+ \leq l_{pq} + l_{s\perp} \leq (n-2)\sqrt{2}.$$

The shortest distance from a point on Q_0 to a point on Q_c is at most l_+. Hence, Q_0, Q_c, G_+ form a configuration of $\alpha_{4.2}$ with $L \leq l_+ \leq (n-2)\sqrt{2}$ and with the same container size as the original one.

For case (c), refer to Figure 30. Define $i_{0c}, i_{c0}, i_{c2}, G_{pq}, G_s$ similarly. Additionally, let i_{c1} be the intersection point of Q_c and its next square in K_1. Draw line segment, denoted G_r with length l_r, from i_{c1} to r.

Counting the number of squares,

$$\left\lceil \frac{l_{pq}}{\sqrt{2}} \right\rceil + \left\lceil \frac{l_r}{\sqrt{2}} \right\rceil + \left\lceil \frac{l_s}{\sqrt{2}} \right\rceil + 2 \leq n,$$

so

$$l_{pq} + l_r + l_s \leq (n-2)\sqrt{2}.$$

Perform these steps to form an n-configuration of $\alpha_{4.2}$ with the same container size:

(1) Let $i_{c1\perp}$ be a point on Q_c with shortest distance to e_r. Also let $i_{c2\perp}$ be a point on Q_c with shortest distance to e_s. Draw a line segment, denoted $G_{r\perp}$ with length $l_{r\perp}$, perpendicular to e_r, from $i_{c1\perp}$ to e_r. Similarly draw a line segment, denoted $G_{s\perp}$ with length $l_{s\perp}$, perpendicular to e_s, from $i_{c2\perp}$ to e_s.
  Note that

$$l_{pq} + l_{r\perp} + l_{s\perp} \leq (n-2)\sqrt{2}.$$

(2) Swap the positions of $G_{s\perp}$ and the whole of Q_c and $G_{r\perp}$, i.e. translate $G_{s\perp}$ such that $G_{s\perp}$ attaches to the end of G_{pq}, and translate Q_c, $G_{r\perp}$ and i_{c0} together such that the end point of $G_{s\perp}$ is point i_{c0}.
   Then Q_c touches e_s.
(3) Swap the positions of $G_{r\perp}$ and Q_c, i.e. translate $G_{r\perp}$ such that $G_{r\perp}$ attaches to the end of $G_{s\perp}$, and translate Q_c and i_{c0} together such that the end of $G_{r\perp}$ is point i_{c0}.
   Now Q_c touches both e_s and e_r.
(4) Draw a line segment, denoted G_+ with length l_+, from i_{0c} to i_{c0}.
   By the triangle inequality,

$$l_+ \leq l_{pq} + l_{r\perp} + l_{s\perp} \leq (n-2)\sqrt{2}.$$

Hence, Q_0, Q_c, G_+ form a configuration of $\alpha_{4.2}$ with the same container size as the original one.

□

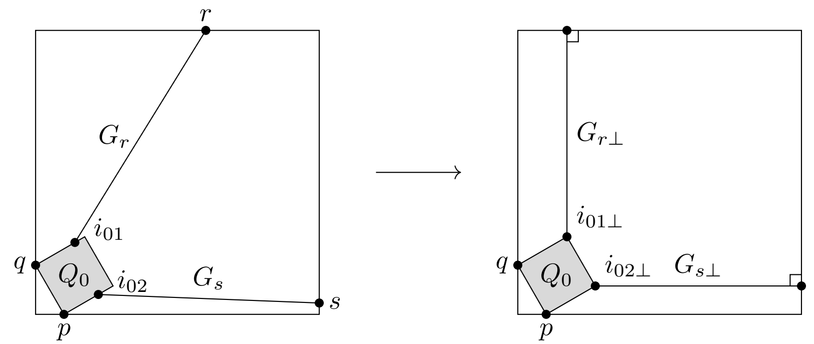


FIGURE 28. Transformations made to case (a), showing the diagram before and after.

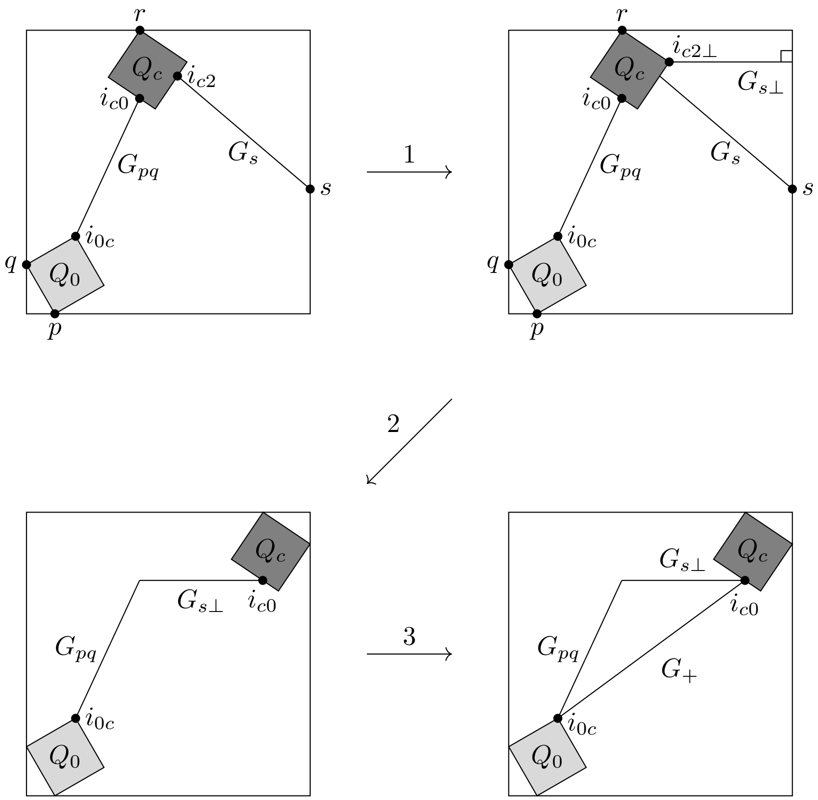


FIGURE 29. Transformations made to case (b), showing the diagram before and after each step.

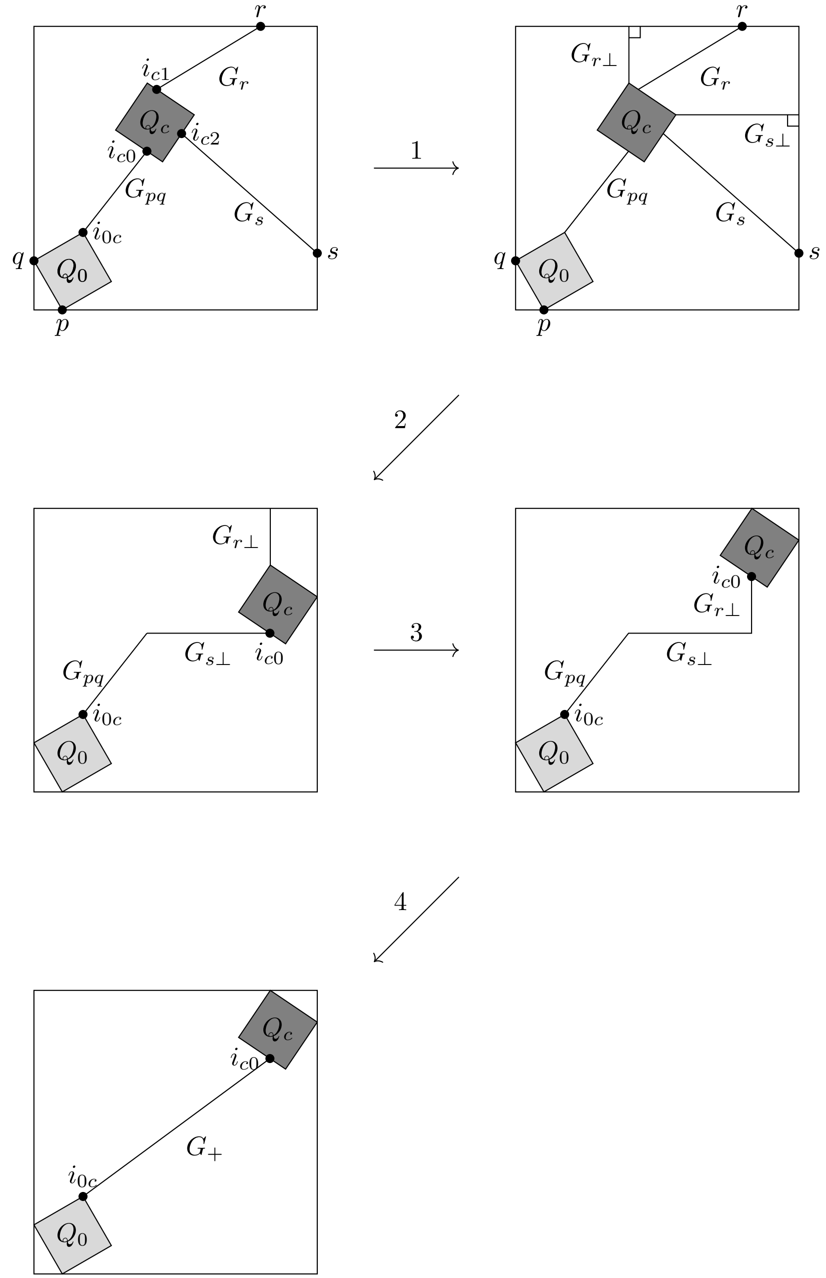


Figure 30. Transformations made to case (c), showing the diagram before and after each step.

A similar method is used in proving class $\alpha_{4.4}$. Due to the huge number of cases, only figures will be provided, which should be sufficient for understanding.

Lemma 6.7. *The side length of the square container of any n-configuration of* $\alpha_{4.4}$ *is always smaller than that of* $\alpha_{4.2}$.

Proof. Define p, q, r, s and e_p, e_q, e_r, e_s the same way as in Lemma 6.6, with the additional assumption that e_p is opposite to e_r and that e_q is opposite to e_s. Let Q_p, Q_q, Q_r, Q_s be the distinct unit squares that p, q, r, s are on respectively.

Since all squares in a configuration are connected, let K_r be the finite sequence of all distinct squares on a path starting from Q_p and ending at Q_r, including Q_p and Q_r. Similarly let K_q be the sequence from Q_p to Q_q, and let K_s be the sequence from Q_p to Q_s.

Let Q_{cq} be the last square in K_q that is also in K_r, and let Q_{cs} be the last square in K_s that is also in K_r. Q_{cq} and Q_{cs} always exist since Q_p is in all K_q, K_r and K_s.

These are the cases for Q_{cq} and Q_{cs}:

1. $Q_{cq} = Q_p$. Then
   - (a) $Q_{cs} = Q_p$;
   - (b) $Q_{cs} \neq Q_p$ and $Q_{cs} \neq Q_r$; or
   - (c) $Q_{cs} = Q_r$.
2. $Q_{cq} = Q_r$, which is equivalent to case 1 by swapping the roles of p and r.
3. $Q_{cq} \neq Q_p$ and $Q_{cq} \neq Q_r$. Then
   - (a) $Q_{cs} = Q_p$, which is equivalent to case 1(b) by swapping q and s;
   - (b) $Q_{cs} \neq Q_p$ and Q_{cs} is before Q_{cq} in K_r;
   - (c) $Q_{cs} = Q_{cq}$;
   - (d) $Q_{cs} \neq Q_r$ and Q_{cs} is after Q_{cq} in K_r, which is equivalent to case 3(b) by swapping p and r; or
   - (e) $Q_{cs} = Q_r$, which is equivalent to case 1(b) by swapping p and r and also swapping q and s.

Using the same technique in Lemma 6.6, Figures 31 to 35 show the steps to transform the configuration into class $\alpha^{\star}_{4.3}$ with $l_1 + l_2 \leq (n-1)\sqrt{2}$ or class $\alpha_{4.2}$ with $L \leq (n-2)\sqrt{2}$.

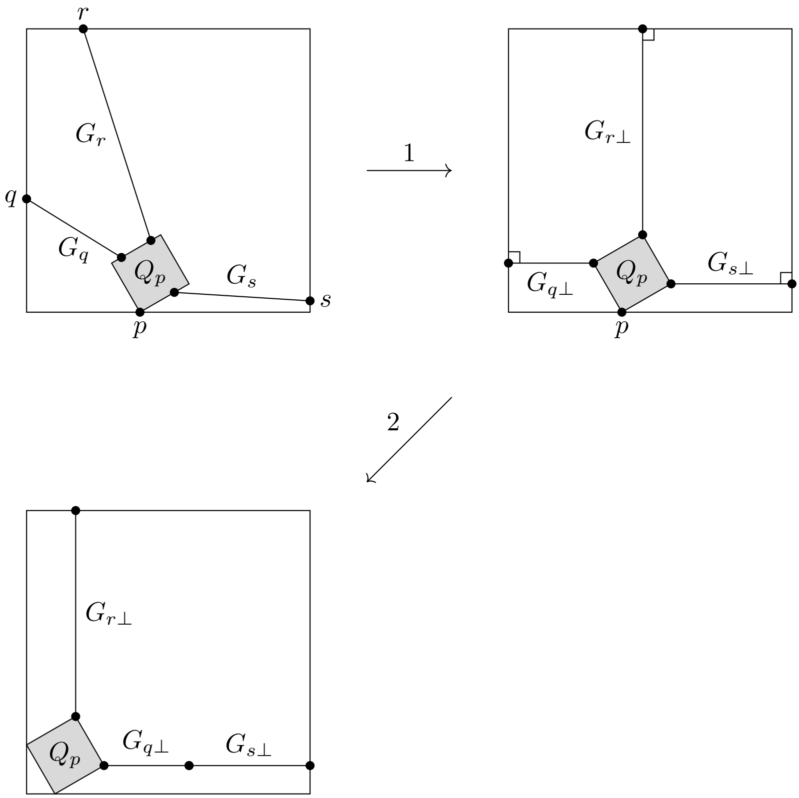


FIGURE 31. Transformation of case 1(a) to class $\alpha^{\star}_{4.3}$.

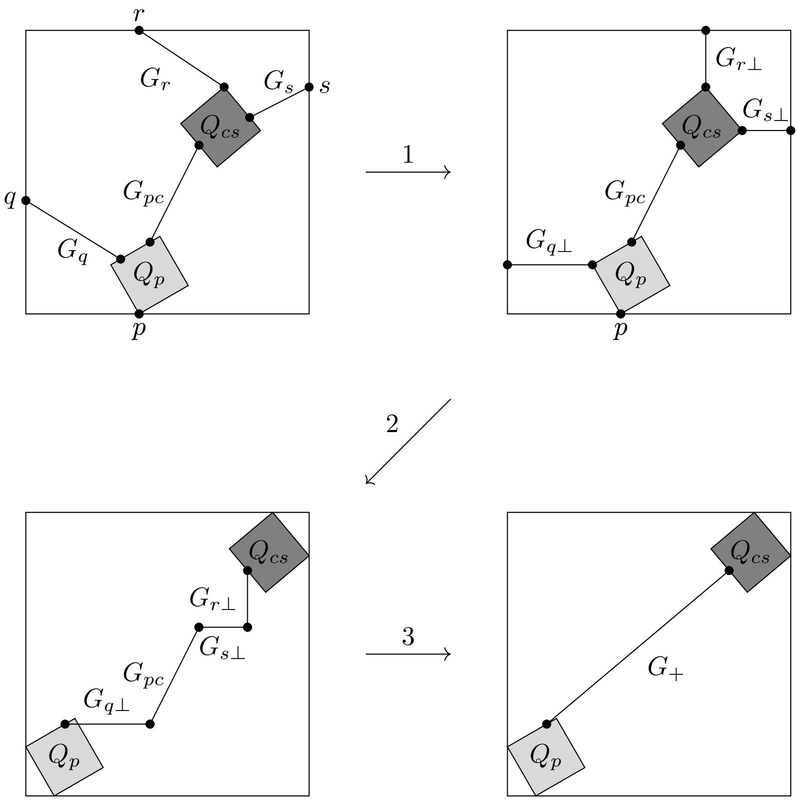


FIGURE 32. Transformation of case 1(b) to class $\alpha_{4.2}$.

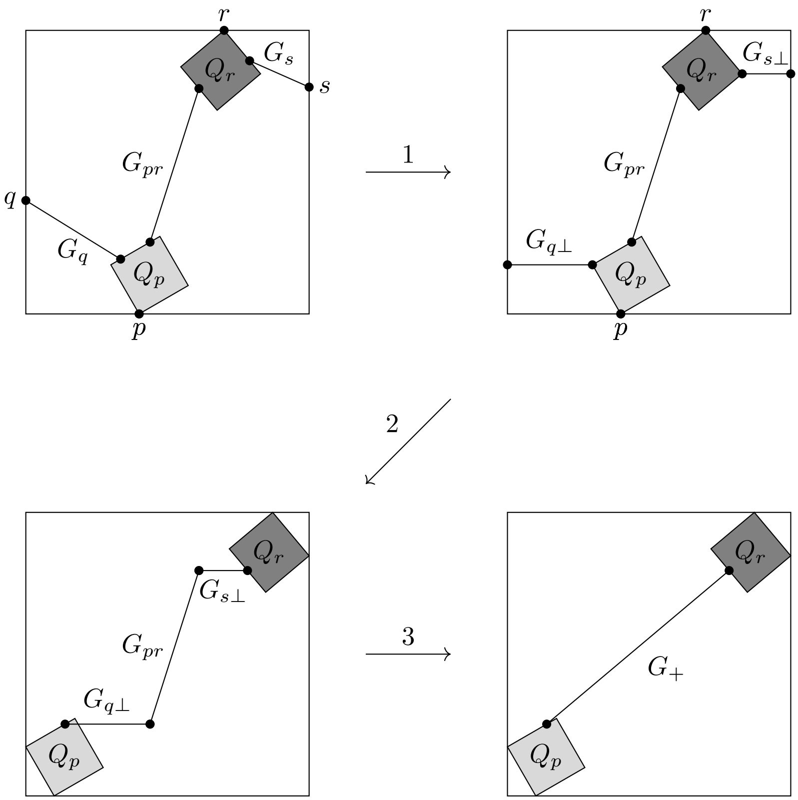


FIGURE 33. Transformation of case 1(c) to class $\alpha_{4.2}$.

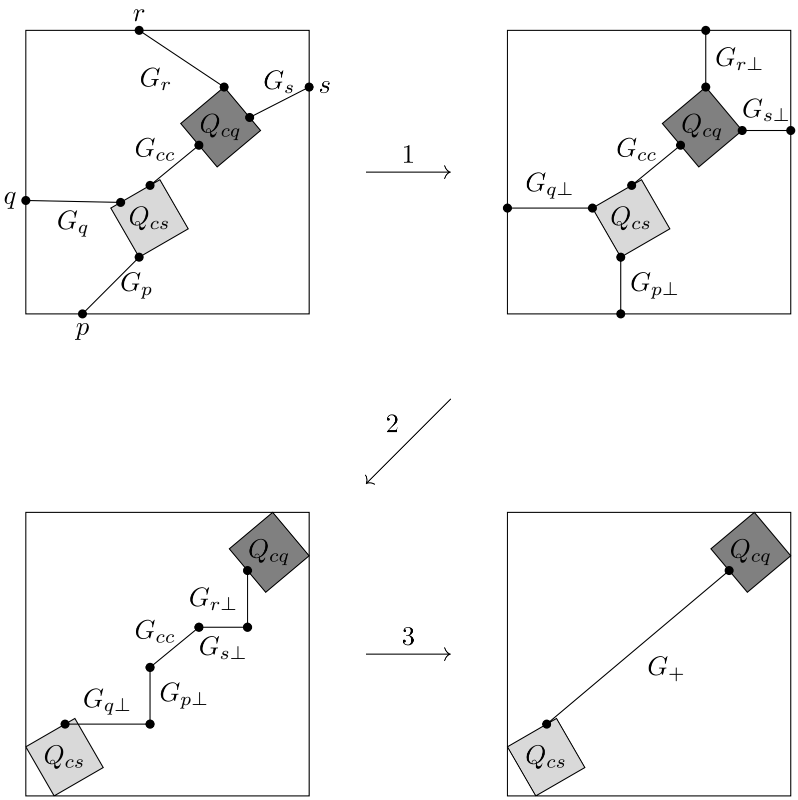


FIGURE 34. Transformation of case 3(b) to class $\alpha_{4.2}$.

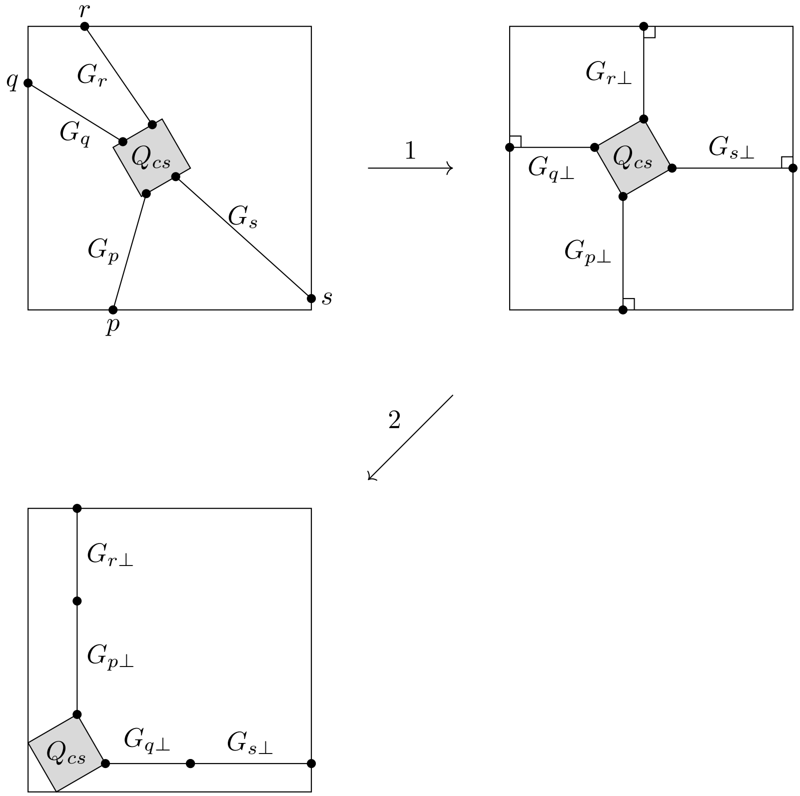


FIGURE 35. Transformation of case 3(c) to class $\alpha^{\star}_{4.3}$.

□

6.5. **Final Step.** Finally, we will relate our previous results to finding $s(n)$. Let us summarise them in the following corollary.

Corollary 6.8. *For $n \geq 2$, if an n-configuration λ and its container $\mathcal{C}$ are of class α_4, then the maximum side length of $\mathcal{C}$ is*

$$n + \sqrt{5} - 2.$$

Proof. By Corollary 5.5, λ and $\mathcal{C}$ are of class $\alpha_{4.2}$, $\alpha_{4.3}$ or $\alpha_{4.4}$. If $\mathcal{C}$ is of class $\alpha_{4.3}$ and $\alpha_{4.4}$, then $\mathcal{C}$ is, by Lemma 6.6 and Lemma 6.7 respectively, smaller than a container of class $\alpha_{4.2}$.

If $\mathcal{C}$ is of class $\alpha_{4.2}$, by Lemma 5.7, λ and $\mathcal{C}$ are of class $\alpha_{4.2}$ type I, II or III. By Lemma 6.2, 6.3 and 6.4, the maximum side length of $\mathcal{C}$ is $(n + \sqrt{5} - 2)$. □

Combining our lemma regarding choice function, we can derive the upper bound for $s(n)$.

Corollary 6.9. *For all integer $n \geq 3$,*

$$s(n) \leq n + \sqrt{5} - 2.$$

Proof. By Lemma 2.8, define $\hat{r}_c : \Lambda_n \to \mathbb{S}^1$ such that $d_\lambda(\hat{r}_c(\lambda)) = d_\lambda(\hat{r}_c(\lambda)_\perp)$ for all $\lambda \in \Lambda_n$. Then by Corollary 6.8,

$$f_\lambda(\hat{r}_c(\lambda)) \leq n + \sqrt{5} - 2$$

for all $\lambda \in \Lambda_n$. Therefore by Lemma 2.13,

$$s(n) \leq n + \sqrt{5} - 2.$$

□

Before we prove a lower bound for $s(n)$ and conclude our proof, we first show a result regarding local maxima of $d_\lambda(\hat{r})$, which will be made use of.

Lemma 6.10. *For some $\vec{v} \in \mathbb{R}^2$, let*

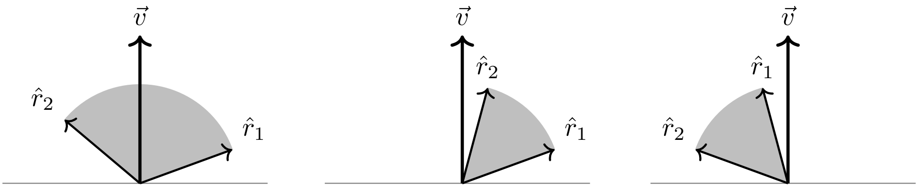
$$d(\hat{r}) = \vec{v} \cdot \hat{r} > 0$$

be a function for $\hat{r} \in \mathbb{S}^1$ from $\hat{r}_1$ going anticlockwise to $\hat{r}_2$. Then $\vec{v} \times \hat{r}_1$ and $\vec{v} \times \hat{r}_2$ have the same direction if and only if $d(\hat{r})$ does not attain local maximum from $\hat{r} = \hat{r}_1$ to $\hat{r}_2$.

Proof. $d(\hat{r})$ attains local maximum at $\hat{r} = \hat{r}_0$ if and only if $\hat{r}_0$ aligns with $\vec{v}$. Since $d(\hat{r}) > 0$, $\vec{v} \cdot \hat{r}_1 > 0$ and $\vec{v} \cdot \hat{r}_2 > 0$. By considering the directions of $\hat{r}_1$ and $\hat{r}_2$ relative to $\vec{v}$, we can determine the directions of the cross products $\vec{v} \times \hat{r}_1$ and $\vec{v} \times \hat{r}_2$, as shown in Table 7.

	Relative directions to $\vec{v}$ of $\hat{r}_1$	Relative directions to $\vec{v}$ of $\hat{r}_2$	Directions of $\vec{v} \times \hat{r}_1$ and $\vec{v} \times \hat{r}_2$	Is $\vec{v}$ between $\hat{r}_1$ and $\hat{r}_2$?
(a)	Clockwise	Anticlockwise	Opposite	Yes
(b)	Clockwise	Clockwise	Same	No
(c)	Anticlockwise	Anticlockwise	Same	No
(d)	Anticlockwise	Clockwise	Impossible, as going from $\hat{r}_1$ to $\hat{r}_2$ is anticlockwise.	

TABLE 7. Relationship between the directions of $\hat{r}_1$ and $\hat{r}_2$ relative to $\vec{v}$, directions of cross products $\vec{v} \times \hat{r}_1$ and $\vec{v} \times \hat{r}_2$, and whether a local maximum is attained.



(A) Case (a) (B) Case (b) (C) Case (c)

FIGURE 36. Cases of the relative directions of $\hat{r}_1$ and $\hat{r}_2$ in Table 7.

□

We have finally reached our main result, which proves the exact value of $s(n)$. Similar to the proof of $s(2)$, we will consider a particular n-configuration and find its f_λ^{min}, which serves as a lower bound for $s(n)$.

Theorem 6.11. *For all integer $n \geq 2$,*

$$s(n) = n + \sqrt{5} - 2.$$

Proof. The case when $n = 2$ is proven in Theorem 3.3. It suffices to show for any $n \geq 3$, an n-configuration λ has $f_\lambda^{min} = n + \sqrt{5} - 2$, as this implies $s(n) \geq n + \sqrt{5} - 2$ and combining with Corollary 6.9, an exact value of $s(n)$ is found.

Consider the n-configuration λ in Figure 37 with its convex hull shown in dotted line. All the coordinates are shown in Table 8.

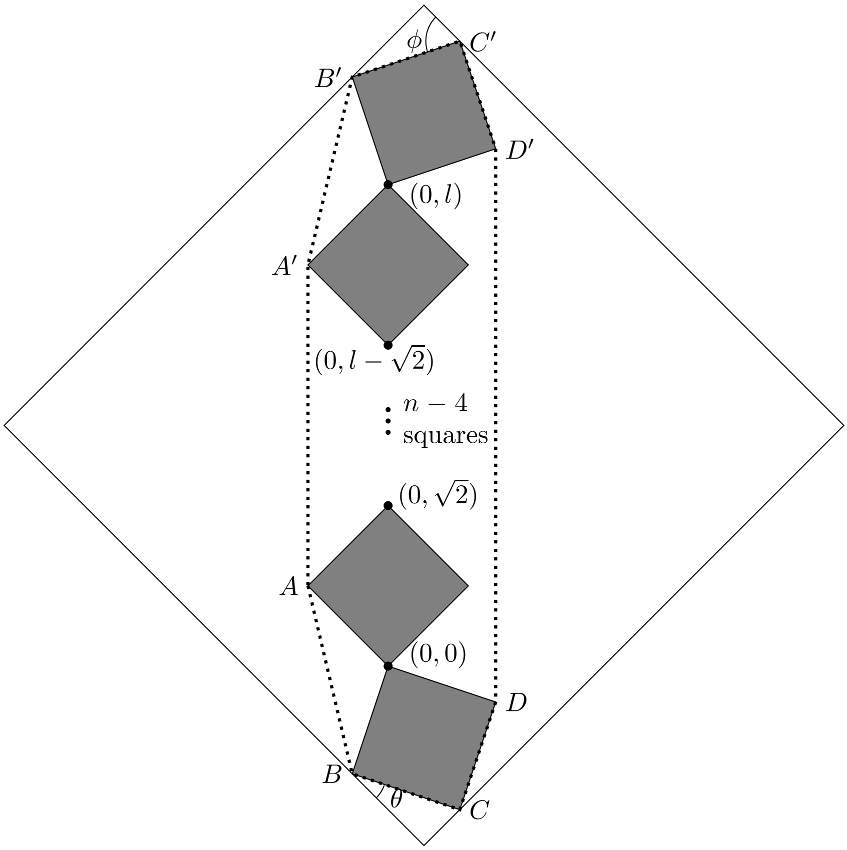


FIGURE 37. The least optimal n-configuration. $\theta = \tan^{-1}\frac{1}{2}$, $\phi = \tan^{-1} 2$, $l = (n-2)\sqrt{2}$.

#	Vertices			
1	$(0,0)$	$B(-\frac{\sqrt{10}}{10}, -\frac{3\sqrt{10}}{10})$	$C(\frac{\sqrt{10}}{5}, -\frac{2\sqrt{10}}{5})$	$D(\frac{3\sqrt{10}}{10}, -\frac{\sqrt{10}}{10})$
2	$(0,0)$	$A(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$	$(0, \sqrt{2})$	$(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$
3	$(0, \sqrt{2})$	$(-\frac{\sqrt{2}}{2}, \frac{3\sqrt{2}}{2})$	$(0, 2\sqrt{2})$	$(\frac{\sqrt{2}}{2}, \frac{3\sqrt{2}}{2})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$(0, (k-2)\sqrt{2})$	$(-\frac{\sqrt{2}}{2}, \frac{(2k-3)\sqrt{2}}{2})$	$(0, (k-1)\sqrt{2})$	$(\frac{\sqrt{2}}{2}, \frac{(2k-3)\sqrt{2}}{2})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	$(0, l-\sqrt{2})$	$A'(-\frac{\sqrt{2}}{2}, l - \frac{\sqrt{2}}{2})$	$(0, l)$	$(\frac{\sqrt{2}}{2}, l - \frac{\sqrt{2}}{2})$
n	$(0, l)$	$B'(-\frac{\sqrt{10}}{10}, l + \frac{3\sqrt{10}}{10})$	$C'(\frac{\sqrt{10}}{5}, l + \frac{2\sqrt{10}}{5})$	$D'(\frac{3\sqrt{10}}{10}, l + \frac{\sqrt{10}}{10})$

TABLE 8. The coordinates of vertices of each square. $l = (n-2)\sqrt{2}$.

Let $\hat{r} = (\cos\gamma, \sin\gamma)$. Notice due to symmetry,

$$d_\lambda\big((\cos\gamma, \sin\gamma)\big) = d_\lambda\big((\cos(\pi-\gamma), \sin(\pi-\gamma))\big), \tag{3}$$

so we now only consider when $0 \le \gamma \le \pi/2$.

We will now show that $d_\lambda(\hat{r})$ is increasing going anticlockwise from $\hat{r} = (1,0)$ to $(0,1)$. Comparing the slopes of AB, $C'D'$ and BC,

$$\text{Slope of } AB = -2-\sqrt{5} < \text{Slope of } C'D' = -3 < \text{Slope of } BC = -\frac{1}{3},$$

so when $\hat{r}$ rotates anticlockwise from $(1,0)$ to $(0,1)$,

- $d_\lambda(\hat{r}) = \overrightarrow{AD'} \cdot \hat{r}$ from $\hat{r} = (1,0)$ up to $\frac{1}{\sqrt{10+4\sqrt{5}}}(2+\sqrt{5}, 1)$; then
- $d_\lambda(\hat{r}) = \overrightarrow{BD'} \cdot \hat{r}$ up to $\hat{r} = \frac{1}{\sqrt{10}}(3,1)$; then
- $d_\lambda(\hat{r}) = \overrightarrow{BC'} \cdot \hat{r}$ up to $\hat{r} = \frac{1}{\sqrt{10}}(1,3)$; then finally
- $d_\lambda(\hat{r}) = \overrightarrow{CC'} \cdot \hat{r}$ up to $\hat{r} = (0,1)$.

Recall that since $n \ge 3$, $l = (n-2)\sqrt{2} \ge \sqrt{2}$. We now verify the directions of these cross products:

$$\overrightarrow{AD'} \times (1,0) = (-l - \frac{1}{\sqrt{10}} + \frac{1}{\sqrt{2}})\hat{k}$$

$$\overrightarrow{AD'} \times (2+\sqrt{5},1) = (-(2+\sqrt{5})l + \sqrt{2} + \frac{3\sqrt{15}}{5})\hat{k}$$

$$\overrightarrow{BD'} \times (2+\sqrt{5},1) = (-(2+\sqrt{5})l - 2\sqrt{2} + \frac{2\sqrt{10}}{5})\hat{k}$$

$$\overrightarrow{BD'} \times (3,1) = (-3l - \frac{4\sqrt{10}}{5})\hat{k}$$

$$\overrightarrow{BC'} \times (3,1) = (-3l - \frac{9\sqrt{10}}{5})\hat{k}$$

$$\overrightarrow{BC'} \times (1,3) = (-l + \frac{\sqrt{10}}{5})\hat{k}$$

Note that all of these cross products can be expressed in $r\hat{k}$ where r is a negative constant i.e. they all have the same direction. So by Lemma 6.10, $d_\lambda(\hat{r})$ has no local maximum from $\hat{r} = (1,0)$ to $\frac{1}{\sqrt{10}}(1,3)$. In addition, $d_\lambda(\hat{r})$ attains a local maximum at $\hat{r} = (0,1)$ since $(0,1)$ is along CC'. Besides, since

$$\frac{d}{d\gamma}\big(\overrightarrow{AD'} \cdot (\cos\gamma, \sin\gamma)\big)\Big|_{\gamma=0} = \overrightarrow{AD'} \cdot (0,1) > 0,$$

$d_\lambda(\hat{r})$ is increasing at $\gamma = 0$. Therefore, $d_\lambda(\hat{r})$ is increasing throughout from $\hat{r} = (1,0)$ to $(0,1)$.

Notice that when $\gamma = \pi/4$, $\hat{r} = \frac{1}{\sqrt{2}}(1,1)$, and

$$d_\lambda(\hat{r}) = \overrightarrow{BC'} \cdot \frac{1}{\sqrt{2}}(1,1) = \sqrt{5} + \frac{l}{\sqrt{2}} = n + \sqrt{5} - 2.$$

so

$$d_\lambda(\hat{r}) < n + \sqrt{5} - 2 \qquad \text{for } \gamma \in \left[0, \frac{\pi}{4}\right),$$
$$d_\lambda(\hat{r}) > n + \sqrt{5} - 2 \qquad \text{for } \gamma \in \left(\frac{\pi}{4}, \frac{\pi}{2}\right].$$

Using Equation (3), we can conclude that

$$d_\lambda(\hat{r}) < n + \sqrt{5} - 2 \qquad \text{for } \gamma \in \left[0, \frac{\pi}{4}\right),$$
$$d_\lambda(\hat{r}) > n + \sqrt{5} - 2 \qquad \text{for } \gamma \in \left(\frac{\pi}{4}, \frac{3\pi}{4}\right),$$
$$d_\lambda(\hat{r}) < n + \sqrt{5} - 2 \qquad \text{for } \gamma \in \left(\frac{3\pi}{4}, \pi\right],$$

so $f_\lambda(\hat{r}) = \max\{d_\lambda(\hat{r}), d_\lambda(\hat{r}_\perp)\} \geq n + \sqrt{5} - 2$ where equality occurs when $\gamma = \pi/4$. Thus,

$$f_\lambda^{min} = n + \sqrt{5} - 2,$$

completing the proof. □

7. Conclusion

In conclusion, the least optimal n-configuration is of class $\alpha_{4.2}$ type I with $\theta = \tan^{-1}\frac{1}{2}$ and $\phi = \tan^{-1} 2$, which gives $s(n) = n + \sqrt{5} - 2$. To reiterate, we use the idea of a rotating container to confirm the existence of a 2-configuration that requires a square container with side length $\sqrt{5}$. Then, we use a square constructing method and Lemma 2.13 that allow us to choose a specific orientation of the container, with some algebra and calculus, to prove that $s(2) = \sqrt{5}$. Finally, we study different ways of how the configuration intersects with the container and prove the final result $s(n) = n + \sqrt{5} - 2$.

Given that the most optimal packing of n unit squares is still an open problem for many values of n, it is surprising that the least optimal square packing problem has an exact and simple formula as the solution.

We believe that this problem can be extended to other shapes such as rectangles with a certain ratio, triangles or to higher dimensions such as cubes or tetrahedrons with the setting of being least optimal. Some techniques harnessed in this research can be made useful to study the packing of other shapes.

At last, although in the least optimal n-configuration, the unit squares turn out to connect to each other vertex-to-vertex, the problem of connecting vertex-to-vertex remains unsolved. Whether there is a more fundamental reason why vertex-to-vertex connection is preferred is unknown. In our research, we mainly investigate in the case where we required the square to intersect with the convex hull on a particular edge, and we found out that the square may not intersect with the convex hull at the ends of the edge to attain minimum container size. However, it is uncertain whether allowing the square to intersect on any edge of an actual configuration will result in a vertex-to-vertex connection attaining minimum container size. The problem of least optimal square packing still requires future work to be done.

References

1. Erich Friedman, *Packing unit squares in squares: A survey and new results*, The Electronic Journal of Combinatorics (2005).
2. Walter Rudin, *Principles of mathematical analysis*, 3rd. ed., McGraw-Hill, 1976.
3. Y. M. Yaglom, D. O. Shklyarsky, and N. N. Chentsov, *Selected problems and theorems from elementary mathematics, part 2 (planimetry)*, 1952.

8. Appendix

8.1. **C++ code for** $s(2)$**.**

```
#include <bits/stdc++.h>
using namespace std;
typedef long long ll;

double dot(pair<double, double> a, pair<double, double> b)
{
    return a.first * b.first + a.second * b.second;
}

int main()
{
    ios_base::sync_with_stdio(0);
    cin.tie(0);
    cout.tie(0);

    setprecision(30);
    cout << fixed;

    double s = -10000;
    int goal_n = 2;
    int scale_y = 10, scale_theta = 1000, scale_angle_of_r = 1000;
    int max_y = 10, max_theta = 1570, max_angle_of_r = 1570;
    int min_y = 0, min_theta = 0, min_angle_of_r = 0;
    int number_of_pts = 8;

    for (int aug_y = min_y; aug_y <= max_y; ++aug_y)
    {
        for (int aug_t = min_theta; aug_t <= max_theta; ++aug_t)
        {

            double t = (double)aug_t / scale_theta;
            double y = (double)aug_y / scale_y;

            vector<pair<double, double>> pts;
            pts.push_back({0, 0});
            pts.push_back({1, 0});
            pts.push_back({0, 1});
            pts.push_back({1, 1});
            pts.push_back({1, y});
            pts.push_back({1 + sin(t), y - cos(t)});
            pts.push_back({1 + sin(t) + cos(t), y + sin(t) - cos(t)});
            pts.push_back({1 + cos(t), y + sin(t)});

            double f_min = 10000;

            for (int aug_angle_of_r = min_angle_of_r; aug_angle_of_r <= max_angle_of_r; ++aug_angle_of_r)
            {

                double angle_of_r = (double)aug_angle_of_r / scale_angle_of_r;
                pair<double, double> r = {cos(angle_of_r), sin(angle_of_r)};
                pair<double, double> r_perp = {sin(angle_of_r), -cos(angle_of_r)};

                vector<double> pts_dot_r;
                vector<double> pts_dot_r_perp;
                for (int i = 0; i < number_of_pts; ++i)
                {
                    pts_dot_r.push_back(dot(pts[i], r));
                    pts_dot_r_perp.push_back(dot(pts[i], r_perp));
                }

                double min_pts_dot_r = 10000, min_pts_dot_r_perp = 10000;
                double max_pts_dot_r = -10000, max_pts_dot_r_perp = -10000;
                for (int i = 0; i < number_of_pts; ++i)
                {
```

```
                    min_pts_dot_r = fmin(min_pts_dot_r, pts_dot_r[i]);
                    max_pts_dot_r = fmax(max_pts_dot_r, pts_dot_r[i]);
                    min_pts_dot_r_perp = fmin(min_pts_dot_r_perp, pts_dot_r_perp[i]);
                    max_pts_dot_r_perp = fmax(max_pts_dot_r_perp, pts_dot_r_perp[i]);
                }
                double d = fmax(max_pts_dot_r - min_pts_dot_r, max_pts_dot_r_perp - min_pts_dot_r_perp);
                f_min = fmin(f_min, d);
            }
            if (f_min >= goal_n - 2 + 2.2364)
            {
                cout << "y: " << y << " t: " << t << " f_min: " << f_min << endl;
            }
            s = fmax(s, f_min);
        }
    }
    cout << "s = " << s;

    return 0;
}
```

8.2. $s(2)$ **Approximate Results.**

```
y: 0.000000 t: 0.907000 f_min: 2.236402
y: 0.000000 t: 0.909000 f_min: 2.236423
y: 0.000000 t: 0.911000 f_min: 2.236442
y: 0.000000 t: 0.913000 f_min: 2.236459
y: 0.000000 t: 0.915000 f_min: 2.236474
y: 0.000000 t: 0.917000 f_min: 2.236486
y: 0.000000 t: 0.919000 f_min: 2.236497
y: 0.000000 t: 0.921000 f_min: 2.236505
y: 0.000000 t: 0.923000 f_min: 2.236510
y: 0.000000 t: 0.925000 f_min: 2.236514
y: 0.000000 t: 0.927000 f_min: 2.236515
y: 0.000000 t: 0.929000 f_min: 2.236514
y: 0.000000 t: 0.931000 f_min: 2.236511
y: 0.000000 t: 0.933000 f_min: 2.236505
y: 0.000000 t: 0.935000 f_min: 2.236497
y: 0.000000 t: 0.937000 f_min: 2.236487
y: 0.000000 t: 0.939000 f_min: 2.236475
y: 0.000000 t: 0.941000 f_min: 2.236461
y: 0.000000 t: 0.943000 f_min: 2.236444
y: 0.000000 t: 0.945000 f_min: 2.236425
y: 0.000000 t: 0.947000 f_min: 2.236404
y: 1.000000 t: 0.636000 f_min: 2.236408
y: 1.000000 t: 0.638000 f_min: 2.236415
y: 1.000000 t: 0.640000 f_min: 2.236420
y: 1.000000 t: 0.642000 f_min: 2.236423
y: 1.000000 t: 0.644000 f_min: 2.236424
y: 1.000000 t: 0.646000 f_min: 2.236422
y: 1.000000 t: 0.648000 f_min: 2.236419
y: 1.000000 t: 0.650000 f_min: 2.236413
y: 1.000000 t: 0.652000 f_min: 2.236404
s = 2.236515
```

REVIEWERS' COMMENTS

The authors raised a problem that is opposite to the classical square packing problem, and called it "the least optimal square packing problem": given n unit squares in the plane, for each configuration where the squares are connected to each other, can we pack this configuration into a large square container? This problem studies the configuration that requires that largest container.

Reviewers generally think that the article is well-written, and demonstates that the authors have a rather broad understanding of mathematical techniques, ranging from numerical methods to analytical methods.