# Arrangements of Stars on the American Flag

Dimitris Koukoulopoulos and Johann Thiel

**Abstract**

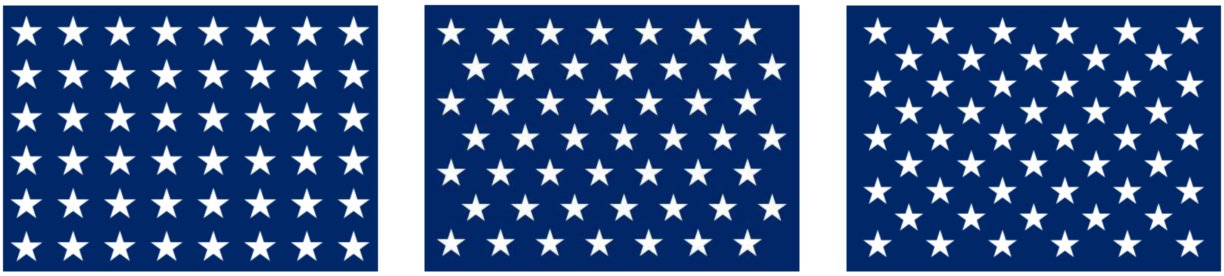
In this article, we examine the existence of nice arrangements of stars on the American flag. We show that despite the existence of such arrangements for any number of stars from 1 to 100, with the exception of 29, 69 and 87, they are rare as the number of stars increases.

## 1 Introduction

The Union Jack is the blue, upper left portion of the American flag that contains a star for each state in the Union[1]. From 1777 to 2002, the Union Jack was the maritime flag used by American Navy vessels to establish their nationality.

With the passage of the Puerto Rico Democracy Act by Congress in 2010, the door has been opened for Puerto Rico to set a referendum which could result in it becoming the 51st state. By law, each time a state is admitted into the Union, the government must create a new Union Jack with an extra star by the fourth of July following the admittance of the new state. In a recent Slate article [8], Wilson considered this possibility and raised the question: What might a 51 star Union Jack look like?

The current 50 star jack design is usually credited to Robert G. Heft [7]. In 1958, when the debate over Alaska's statehood status was being considered, Heft claims to have guessed correctly that Congress would eventually allow Alaska to join the Union only if Hawaii was admitted as well. Back then Alaska was Democrat territory, while Hawaii was more Republican, so admitting both would maintain the current balance of power in Congress. While some designers were working on a 49 star arrangement, Heft had already created a 50 star arrangement one year ahead of Hawaii becoming the 50th state (see Figure 1).



(a) 48 Stars (b) 49 Stars (c) 50 Stars

Figure 1: Union Jacks from 1959-present

In [8], Wilson identifies six Union Jack patterns that exhibit a lot of symmetry. The six patterns are defined below (see Figure 2 for examples):

> **Long** - This pattern is made by alternating long and short rows of stars, where the long rows contain one more star than the short rows. In this pattern the first and last row are both long.

[1]This Union Jack is not to be confused with the Union Flag of the United Kingdom, which is also known as the Union Jack.

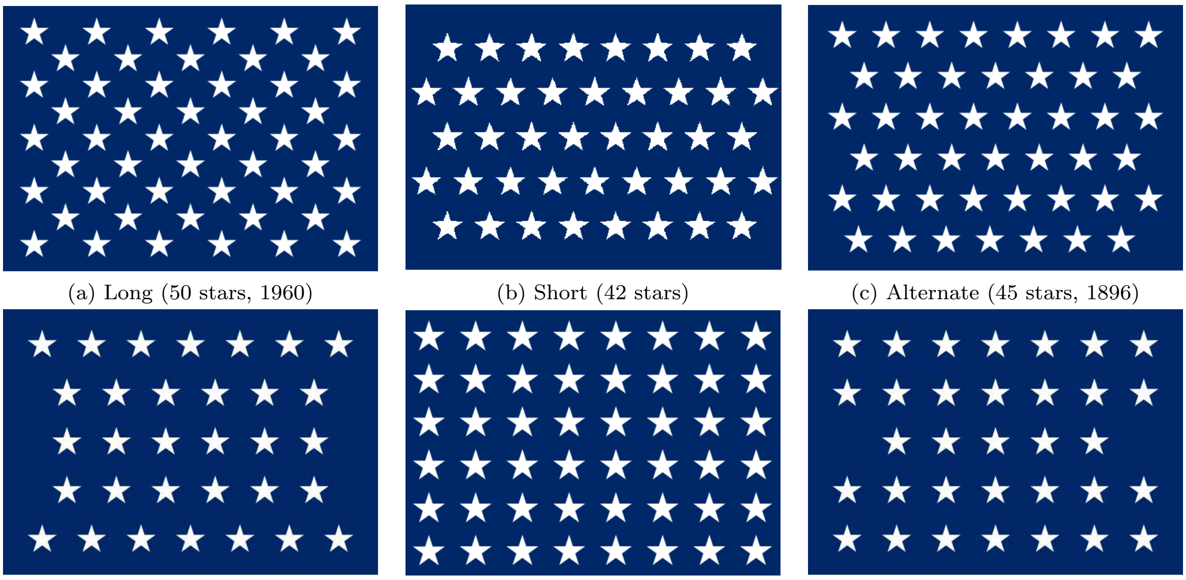
**Short** - This pattern is defined in the same way as the long pattern, except that the first and the last row are both short.

**Alternate** - This pattern is defined in the same way as the long pattern, except that the first row is long and the last row is short, or vice versa.

**Wyoming**[2] - In this pattern, the first and the last row are long, while all other rows are short.

**Equal** - All rows in this pattern have the same number of stars.

**Oregon**[2] - Similar to the equal pattern, all rows in this pattern are of the same length with the exception of the middle row, which is two stars shorter. This pattern requires an odd number of rows.



(a) Long (50 stars, 1960) (b) Short (42 stars) (c) Alternate (45 stars, 1896)

(d) Wyoming (32 stars, 1858) (e) Equal (48 stars, 1912) (f) Oregon (33 stars, 1859)

Figure 2: Union Jack patterns

Implicit in Wilson's definition is the need to have a good ratio of columns to rows to prevent the creation of degenerate Union Jacks (imagine a Union Jack with one row of 50 stars!). For our purposes, we will assume that this ratio is in the interval $[1, 2]$.

**Definition** A *nice* arrangement of stars on the Union Jack is one that uses one of the six patterns defined above with a column to row ratio in the interval $[1, 2]$.

Wilson's article includes a flash program[3], created by Skip Garibaldi, that finds these nice arrangements of $n$ stars on the Union Jack for $n$ from 1 to 100. Using the above definition, Garibaldi's program shows that all integers from 1 to 100, except for 29, 69 and 87, have at least one nice arrangement. It is then natural to ask, for which numbers $n$ do there exist nice arrangements of stars on the Union Jack? To this end, for an integer $N$, we define

$$S(N) = \#\{n \leq N : \text{there is a nice arrangement of } n \text{ stars on the Union Jack}\},$$

so that $S(100) = 97$, by Garibaldi's computation.

[2]These patterns get their names from the state that first caused their introduction.
[3]`http://img.slate.com/media/19/flag.swf`

By analyzing the type of integers counted by $S(N)$ (see Section 2), we will show that the problem of arranging stars on the Union Jack is connected with an old problem due to Erdős about multiplication tables of integers (see Sections 3, 4). This problem has been studied extensively in the literature and the state of the art on it is given by a deep result due to Ford (see Theorem 3.2). Using the available information on the Erdős multiplication table problem, we shall estimate $S(N)$. In particular, we will show that nice arrangements on the Union Jack are actually rare,

$$\lim_{N\to\infty} \frac{S(N)}{N} = 0. \tag{1}$$

## 2 Characterization of nice arrangements

In order to understand when there exists a nice arrangement of stars on the Union Jack for an arbitrary $n$, let us first examine why it is impossible to produce a nice arrangement on the Union Jack with 29 stars. We cannot make a nice 29 star flag with the equal pattern because 29 is prime. For the Wyoming pattern, let $b$ represent the length of the short rows and let $a$ represent the total number of rows. Then there are 2 rows of length $b+1$ and $a-2$ rows of length $b$. Thus, for a nice 29 star flag with the Wyoming pattern, we would need $a$ and $b$ such that $29 = 2(b+1) + (a-2)b = ab + 2$, or $27 = ab$. In addition, our requirement that the column to row ratio must be in the interval $[1, 2]$ leads to the constraint $1 \le (b+1)/a \le 2$. Since 27 cannot be written as a product $ab$ with $1 \le (b+1)/a \le 2$, we cannot use the Wyoming pattern for 29. A similar argument shows that the Oregon pattern fails for 29 because $31 = 29 + 2$ is prime.

The story for the long, short, and alternate patterns requires just a little more work. As an example, suppose we want to create an alternate Union Jack pattern for $n$ stars. Let $b$ equal the number of stars on the long rows and $a$ equal the total number of long rows. Then we should have a representation $n = ba + (b-1)a = 2ab - a = a(2b-1)$, along with the constraint $1 \le b/(2a) \le 2$. For $n = 29$, such a representation is impossible. In the case of the long and short patterns, a similar argument shows that we would need factorizations of $2n-1$ or $2n+1$, respectively, with the factors being close to each other. In the case $n = 29$, we would need one of these restricted factorizations for $57 = 3 \cdot 19$ or 59 (prime).

The discussion on attempts to produce a nice arrangement of 29 stars showed that the problem boils down to searching for very restricted factorizations of 27, 29, 31, 57 or 59. The same arguments lead to the following proposition, which characterizes the values $n$ for which there is a nice arrangement of stars.

**Proposition 2.1** *A nice arrangement of $n$ stars on the Union Jack exists if, and only, if at least one of the following holds:*

*(i) For the long pattern,* $2n-1 = (2a+1)(2b+1)$ *with* $1 \le (b+1)/(2a+1) \le 2$.

*(ii) For the short pattern,* $2n+1 = (2a+1)(2b+1)$ *with* $1 \le (b+1)/(2a+1) \le 2$.

*(iii) For the alternate pattern,* $n = a(2b-1)$ *with* $1 \le b/(2a) \le 2$.

*(iv) For the Wyoming pattern,* $n-2 = ab$ *with* $1 \le (b+1)/a \le 2$.

*(v) For the equal pattern,* $n = ab$ *with* $1 \le b/a \le 2$.

*(vi) For the Oregon pattern,* $n+2 = (2a+1)b$ *with* $1 \le b/(2a+1) \le 2$.

Using the above characterization, it is clear that in order to estimate $S(N)$ we need to understand for how many integers $n \le N$ there exist restricted factorizations of $n$, $n-2$, $n+2$, $2n+1$ or $2n-1$, as given in the proposition above. As we will see, this is related to an old problem of Erdős known as the *multiplication table problem*. We shall discuss this problem in the next section and see how our knowledge of it can be used to bound $S(N)$ in Section 4.

## 3 The Multiplication Table Problem

Consider the $10 \times 10$ multiplication table (see Figure 3a). There are 100 entries in it, but not all are distinct. For example, 12 appears four times in the table, corresponding to the four factorizations $12 = 2 \cdot 6 = 3 \cdot 4 = 4 \cdot 3 = 6 \cdot 2$. If we remove all of the duplicate entries, we see that only 42 numbers remain (see Figure 3b).

| $\times$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 2 | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 |
| 3 | 3 | 6 | 9 | 12 | 15 | 18 | 21 | 24 | 27 | 30 |
| 4 | 4 | 8 | 12 | 16 | 20 | 24 | 28 | 32 | 36 | 40 |
| 5 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
| 6 | 6 | 12 | 18 | 24 | 30 | 36 | 42 | 48 | 54 | 60 |
| 7 | 7 | 14 | 21 | 28 | 35 | 42 | 49 | 56 | 63 | 70 |
| 8 | 8 | 16 | 24 | 32 | 40 | 48 | 56 | 64 | 72 | 80 |
| 9 | 9 | 18 | 27 | 36 | 45 | 54 | 63 | 72 | 81 | 90 |
| 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

(a) All entries

| $\times$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 2 | | | | | | 12 | 14 | 16 | 18 | 20 |
| 3 | | | | | 15 | | 21 | 24 | 27 | 30 |
| 4 | | | | | | | 28 | 32 | 36 | 40 |
| 5 | | | | | 25 | | 35 | | 45 | 50 |
| 6 | | | | | | | 42 | 48 | 54 | 60 |
| 7 | | | | | | | 49 | 56 | 63 | 70 |
| 8 | | | | | | | | 64 | 72 | 80 |
| 9 | | | | | | | | | 81 | 90 |
| 10 | | | | | | | | | | 100 |

(b) Distinct entries

Figure 3: The $10 \times 10$ multiplication table

In 1955, Erdős [2] asked what happens if one considers larger multiplication tables, namely, what is the asymptotic behavior of

$$A(N) = \#\{n \le N : n = m_1 m_2,\ m_1 \le \sqrt{N},\ m_2 \le \sqrt{N}\}$$

as $N$ grows to infinity?

Using a simple argument based on the number of prime factors of an integer, Erdős proved that

$$\lim_{N\to\infty} \frac{A(N)}{N} = 0, \tag{2}$$

that is, most of the integers up to $N$ do not appear as entries in the $\lfloor\sqrt{N}\rfloor \times \lfloor\sqrt{N}\rfloor$ multiplication table. We outline Erdős' argument below.

We start with some remarks about the arithmetic function $\Omega(n)$, which is defined to be the number of prime factors of $n$ counted with multiplicity. Mertens' estimate[4] [1, Theorem 4.12]

$$\sum_{p\le N} \frac{1}{p} = \log\log N + O(1)$$

[4]The variable $p$ always denotes a prime number. Also, log denotes the natural logarithm.

implies that[5]

$$
\begin{aligned}
\frac{1}{N}\sum_{n=1}^{N}\Omega(n) &= \frac{1}{N}\sum_{n\le N}\sum_{\substack{p^a|n\\ a\ge 1}} 1\\
&= \frac{1}{N}\sum_{\substack{p^a\le N\\ a\ge 1}}\sum_{\substack{n\le N\\ p^a|n}} 1\\
&= \frac{1}{N}\sum_{\substack{p^a\le N\\ a\ge 1}}\left(\frac{N}{p^a}+O(1)\right)\\
&= \sum_{p\le N}\frac{1}{p}+O\Bigg(\sum_{\substack{p \text{ prime}\\ a\ge 2}}\frac{1}{p^a}\Bigg)\\
&= \log\log N+O(1).
\end{aligned}
$$

In other words, the average number of prime divisors, counted with multiplicity, of an integer $n$ is about $\log\log n$. Actually, Hardy and Ramanujan proved the following more precise result, which roughly states that, for most positive integers $n$, $\Omega(n)$ is very close to its expected value $\log\log n$.

**Theorem 3.1 (Hardy, Ramanujan, [6])** *For any fixed $\epsilon > 0$ we have that*

$$\lim_{N\to\infty}\frac{1}{N}\#\{n\le N : (1-\epsilon)\log\log N \le \Omega(n) \le (1+\epsilon)\log\log N\} = 1.$$

With the above theorem at our disposal, it is not very hard to show (2). Indeed, if an integer $n$ is counted by $A(N)$, then we can write $n = n_1 n_2$ for some $n_1 \le \sqrt{N}$ and $n_2 \le \sqrt{N}$ such that either

$$\Omega(n_1) \le \frac{2}{3}\log\log N \tag{3}$$

or

$$\Omega(n_2) \ge \Omega(n_1) \ge \frac{2}{3}\log\log N. \tag{4}$$

The number of integers $n \le N$ for which (3) holds is at most

$$\#\Big\{n_1 \le \sqrt{N} : \Omega(n_1) \le \frac{2}{3}\log\log N\Big\}\cdot\#\{n_2\le\sqrt{N}\}.$$

Moreover, if $n$ satisfies (4), then we must have

$$\Omega(n) = \Omega(n_1)+\Omega(n_2) > \frac{4}{3}\log\log N.$$

Hence,

$$\frac{A(N)}{N} \le \frac{1}{\sqrt{N}}\#\Big\{n_1\le\sqrt{N} : \Omega(n_1)\le\frac{2}{3}\log\log N\Big\} + \frac{1}{N}\#\Big\{n\le N : \Omega(n)\ge\frac{4}{3}\log\log N\Big\}.$$

Using Theorem 3.1 to estimate the quantities on the right-hand side of the above inequality, we deduce (2).

[5] The notation $a|b$ means that $a$ divides $b$

The argument given above is quite flexible. By optimizing it and combining it with a stronger version of Theorem 3.1, Erdős [3] showed that for all $\epsilon > 0$ and all $N \geq 3$

$$A(N) \leq \frac{c_0(\epsilon)N}{(\log N)^{\delta-\epsilon}}, \quad \text{where} \quad \delta = 1 - \frac{1+\log\log 2}{\log 2} = 0.086071\ldots \tag{5}$$

and $c_0(\epsilon)$ is a constant that depends only on $\epsilon$. This estimate was subsequently improved by Hall and Tenenbaum [5], who showed that

$$A(N) \leq \frac{c_1 N}{(\log N)^{\delta}\sqrt{\log\log N}} \quad (N \geq 3)$$

for some constant $c_1 > 0$. Both [3] and [5] supplied lower bounds to $A(N)$ as well. The precise size of $A(N)$, up to multiplicative constants, was determined by Ford in [4].

**Theorem 3.2 (Ford, [4])** *There are constants $c_2$ and $c_3$ such that*

$$\frac{c_2 N}{(\log N)^{\delta}(\log\log N)^{3/2}} \leq A(N) \leq \frac{c_3 N}{(\log N)^{\delta}(\log\log N)^{3/2}} \quad (N \geq 3).$$

With Theorem 3.2, we are now able to tackle the question of nice arrangements of stars on the Union Jack.

# 4 Connecting the two problems

Let $S_{equal}(N)$ count the number of positive integers up to $N$ that admit a nice arrangement of stars on the Union Jack using the equal pattern. By Proposition 2.1,

$$\begin{aligned} S_{equal}(N) &= \#\{n \leq N : n = ab \text{ with } 1 \leq b/a \leq 2\} \\ &\leq \#\{n \leq N : n = ab \text{ with } a, b \leq \sqrt{2N}\} \\ &\leq A(2N). \end{aligned}$$

As in Section 2, handling the cases for the long, short and alternate arrangements requires a little more work. Let $S_{long}(N)$ count the number of positive integers up to $N$ that admit a nice arrangement of stars on the Union Jack using the long pattern. Using Proposition 2.1 again,

$$\begin{aligned} S_{long}(N) &= \#\{n \leq N : 2n-1 = (2a+1)(2b+1) \text{ with } 1 \leq (b+1)/(2a+1) \leq 2\} \\ &\leq \#\{n \leq 2N : n = ab \text{ with } 2 \leq (b+1)/a \leq 4\} \\ &\leq \#\{n \leq 2N : n = ab \text{ with } 1 \leq b/a \leq 4\} \\ &\leq \#\{n \leq 2N : n = ab \text{ with } a, b \leq \sqrt{8N}\} \\ &\leq A(8N). \end{aligned}$$

By using similar arguments for $S_{short}(N)$, $S_{alternate}(N)$, $S_{Wyoming}(N)$ and $S_{Oregon}(N)$, defined analogously to $S_{equal}(N)$ and $S_{long}(N)$, we obtain that each of these quantities is bounded by $A(12N)$. Hence,

$$S(N) \leq S_{\text{long}}(N) + S_{\text{short}}(N) + S_{\text{alternate}}(N) + S_{\text{Wyoming}}(N) + S_{\text{equal}}(N) + S_{\text{Oregon}}(N) \leq 6A(12N).$$

Applying Theorem 3.2 to the right-hand side of the above inequality and setting $c_4 = 72c_3$, we deduce our main result.

**Theorem 4.1** *There is a constant $c_4$ such that*

$$S(N) \leq \frac{c_4 N}{(\log N)^{\delta}(\log\log N)^{3/2}} \quad (N \geq 3),$$

*where $\delta$ is defined as in* (5). *In particular, relation* (1) *holds.*

## 5 Concluding Remarks

It is natural to ask if a result analogous to Theorem 3.2 holds for $S(N)$ (i.e., if the upper bound given by Theorem 4.1 is also a lower bound with some other constant in place of $c_4$). This is indeed the case but one has to use a stronger result than Theorem 3.2. Note that

$$S(N) \geq S_{equal}(N) = \#\{n \leq N : \text{there is } d|n \text{ such that } n/2 \leq d^2 \leq n\}.$$

Then we can combine Corollary 2 and Theorem 2 in [4] to show that

$$S(N) \geq S_{equal}(N) \geq \frac{c_5 N}{(\log N)^{\delta}(\log\log N)^{3/2}} \quad (N \geq 3)$$

for some constant $c_5 > 0$.

Several comments on Wilson's online article mentioned one possible way to deal with the lack of a nice 69 star Union Jack by noticing that $69 = 8+7+8+8+7+8+8+7+8$. This row pattern would involve expanding our current definition of a nice arrangement to include patterns whose rows are periodic with period 3. This expanded definition would not alter our final result while allowing more small integers to admit a Union Jack with a nice arrangement. Similarly, other modest expansions to the definition of a nice arrangement of stars on the Union Jack, such as any periodic pattern of long and short rows of any fixed period, or the already considered patterns with an aspect ratio lying in some longer interval $[\alpha, \beta]$, $0 < \alpha < \beta < \infty$, do not affect our result; the set of nice arrangements remains sparse.

Despite the fact that 29 has no nice arrangement of stars, there was a time when a 29 star Union Jack was needed (Figure 4). One can always come up with *some* arrangement of stars for any number but in the long run the U.S. is going to have to get more and more creative to accommodate additional states (and stars!) into the Union.

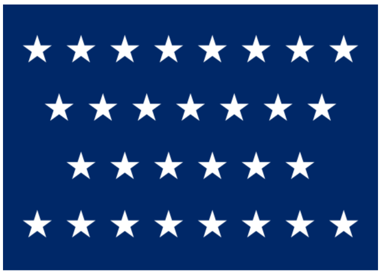


Figure 4: 29 star Union Jack

**Acknowledgments** We would like to thank Sylvia Carlisle for bringing Wilson's article into our attention, and A. J. Hildebrand and Kevin Ford for various helpful suggestions.

## References

[1] Apostol, T. M., Introduction to analytic number theory, Undergraduate Texts in Mathematics, *Springer-Verlag, New York, NY*, 1976.

[2] Erdős, P., Some remarks on number theory, *Riveon Lematematika* **9** (1955), 45–48.

[3] —, An asymptotic inequality in the theory of numbers (Russian), *Vestnik Leningrad. Univ.* **15** (1960), no. 13, 41–49.

[4] Ford, K., The distribution of integers with a divisor in a given interval, *Ann. of Math. (2)* **168** (2008), no. 2, 367–433.

[5] Hall, R. R.; Tenenbaum, G., Divisors, vol. 90, *Cambridge Tracts in Mathematics, Cambridge*, 1988.

[6] Hardy, G. H.; Ramanujan, S., The normal number of prime factors of a number $n$ [Quart. J. Math. **48** (1917), 7692]. *Collected papers of Srinivasa Ramanujan*, 262-275, *AMS Chelsea Publ., Providence, RI*, 2000.

[7] Rasmussen, F. N., A half-century ago, new 50-star American flag debuted in Baltimore (2010), available at `http://articles.baltimoresun.com/2010-07-02/news/bs-md-backstory-1960-flag-20100702_1_48-star-flag-blue-canton-fort-mchenry`.

[8] Wilson, C., 13 Stripes and 51 Stars (2010), available at `http://www.slate.com/id/2256250/`.

**Dimitris Koukoulopoulos** received his B.S. in 2006 from the Aristotle University of Thessaloniki and his Ph.D. in 2010 from the University of Illinois at Urbana-Champaign. He wrote his thesis on generalizations of the multiplication table problem under the direction of Kevin Ford. He is currently a CRM-ISM postdoctoral fellow at Université de Montréal.
*Centre de Recherches Mathématiques, Université de Montréal, P.O. Box 6128, Centre-Ville Station, Montréal, QC H3C 3J7, Canada*
*koukoulo@crm.umontreal.ca*

**Johann Thiel** received his Ph.D. in 2011 from the University of Illinois at Urbana-Champaign under the supervision of A.J. Hildebrand. He is currently an assistant professor at the United States Military Academy.
*Department of Mathematical Sciences, United States Military Academy, 601 Thayer Rd., West Point, NY, USA*
*johann.thiel@usma.edu*