A Combinatorial Approach to the Solitaire Game

David AVIS Antoine DEZA Shmuel ONN

Abstract. The classical game of peg solitaire has uncertain origins, but was certainly popular by the time of Louis XIV, and was described by Leibniz in 1710. One of the classical problems concerning peg solitaire is the feasibility issue. An early tool used to show the infeasibility of various peg games is the *rule of three* [Suremain de Missery 1841]. In the 1960s the description of the *solitaire cone* [Boardman and Conway] provides necessary conditions: valid inequalities over this cone, known as pagoda functions, were used to show the infeasibility of various peg games. In this paper, we recall these necessary conditions and present new developments: the *lattice criterion*, which generalizes the rule of three; and results on the strongest pagoda functions, the facets of the solitaire cone.

1 Introduction and Basic Definitions

1.1 Introduction

Peg solitaire is a peg game for one player which is played on a board containing a number of holes. The most common modern version uses a cross shaped board with 33 holes - see Fig. 1 - although a 37 hole board is common in France. Computer versions of the game now feature a wide variety of shapes, including rectangles and triangles. Initially the central hole is empty, the others contain pegs. If in some row (column resp.) two consecutive pegs are adjacent to an empty hole in the same row (column respectively), we may make a *move* by removing the two pegs and placing one peg in the empty hole. The objective of the game is to make moves until only one peg remains in the central hole. Variations of the original game, in addition to being played on different boards, also consider various alternate starting and finishing configurations.

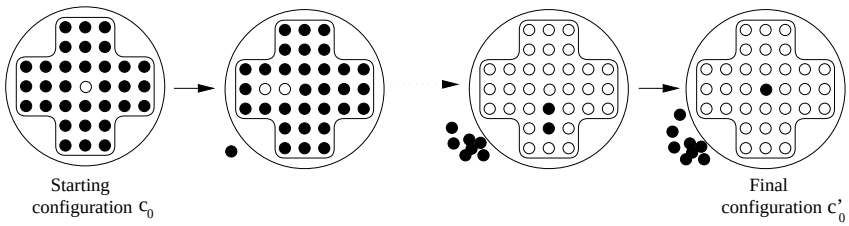

Figure 1: A feasible English solitaire peg game with possible first and last moves

The game itself has uncertain origins, and different legends attest to its discovery by various cultures. An authoritative account with a long annotated bibliography can be found in the comprehensive book of Beasley [3]. The book mentions an engraving of Berey, dated 1697, of a lady with a solitaire board. The book also contains a quotation of Leibniz [6] which was written for the Berlin Academy in 1710. Apparently the first theoretical study of the game that was

published was done in 1841 by Suremain de Missery, and was reported in a paper by Vallot [8]. The modern mathematical study of the game dates to the 1960s at Cambridge University. The group was led by Conway who has written a chapter in [4] on various mathematical aspects of the subject. One of the problems studied by the Cambridge group is the following basic *feasibility* problem (see Definition 1.1 in the sequel for a formal definition):

Peg solitaire feasibility problem. Given a board B and a pair of configurations (c, c') on B, determine if (c, c') is *feasible*, that is, if there is a legal sequence of moves transforming c into c'.

The complexity of the feasibility problem for the game played on a n by n board was shown by Uehara and Iwata [7] to be NP-complete, so easily checked necessary and sufficient conditions for feasibility are unlikely to exist. In this paper, we recall constructions used to prove the infeasibility of some pair (c, c'): the *rule-of-three* in see Section 2, the *solitaire cone* in Section 3.1; and present new developments in Section 3.2 and Section 4.

1.2 Basic definitions

In this section we introduce some terminology used throughout this paper. The *board* of a peg solitaire game is a *finite* subset $B \subset \mathbb{Z}^2$. Thus, B stands for the set of locations (i, j) of holes of the board on which the game is played. For example, the classical 33-board is: $B = \{(i,j) : \ -1 \leq i \leq 1, \ -3 \leq j \leq 3\} \cup \{(i,j) : \ -3 \leq i \leq 3, \ -1 \leq j \leq 1\}$. A *configuration* c on the board is an integer vector $c \in \mathbb{Z}^B \subset \mathbb{R}^B$. It can be interpreted as a configuration of pegs on the board: in the usual game, all configurations c lie in $\{0,1\}^B$, with the interpretation that hole $(i, j) \in B$ contains a peg if $c_{i,j} = 1$ and is empty if $c_{i,j} = 0$; extending this, we allow any integer (possibly negative) number $c_{i,j}$ of pegs to occupy any hole $(i, j) \in B$. The *complement* of a $\{0,1\}$-configuration $c \in \{0,1\}^B$ is defined to be the configuration $\bar{c} := \mathbb{1} - c$ where $\mathbb{1} = (1, 1, \ldots, 1) \in \mathbb{R}^B$ is the all-ones configuration. A *move* or a *jump* μ is a vector in $\mathbb{R}^B$ which has 3 non-zero entries: two entries of -1 in the positions from which pegs are removed and one entry of 1 for the hole receiving the new peg. We can now make the peg solitaire feasibility problem precise.

Definition 1.1 Given a board B and an associated set of moves $\mathcal{M}$, a pair (c, c') of configurations is *feasible* if there is a sequence $\{\mu^1, \ldots, \mu^k\} \subset \mathcal{M}$ of moves on B such that

$$c' - c = \sum_{i=1}^{k} \mu^i : c + \sum_{j=1}^{i} \mu^j \in \{0,1\}^B \text{ for } i = 1, \ldots, k$$

For instance, the English 33-board admits 77 moves; see Fig. 1 for possible first and last moves in some sequence of moves transforming the initial configuration c_0 to its complementary c'_0.

2 The Rule-of-Three

In this section we recall the so-called rule-of-three (cf. [3, 4]), a classical construction used to test solitaire game feasibility. The rule-of-three can be used, for example, to show that on the cross shaped English 33-board, starting with the initial configuration c_0 of Fig. 1, the only reachable final configurations with *exactly one* peg are c'_0 of Fig. 1 and c'_1, c'_2, c'_3 and c'_4 of Fig. 2.
Let $\mathbb{Z}_2 := \{a, b, c, e\}$ be the Abelian group with identity e and addition table $a + a = b + b = c + c = e, \ \ a + b = c, \ a + c = b, \ b + c = a$. Define the following two maps $g_1, g_2 : \mathbb{Z}^2 \longrightarrow \mathbb{Z}_2$,

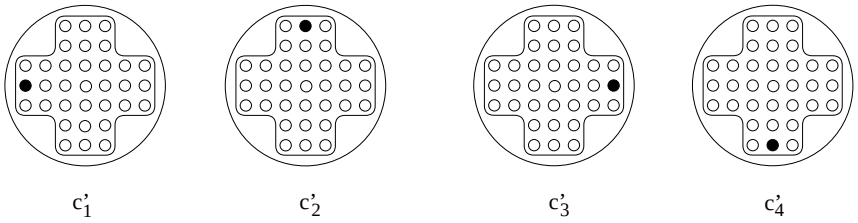

Figure 2: Four other feasible final configurations

which simply color the integer lattice $\mathbb{Z}^2$ by diagonals of a, b and c in either direction; see Fig. 3:

$$g_1(i,j) := \begin{cases} a & \text{if } (i+j) \equiv 0 \pmod 3 \\ b & \text{if } (i+j) \equiv 1 \pmod 3 \\ c & \text{if } (i+j) \equiv 2 \pmod 3 \end{cases}, \qquad g_2(i,j) := \begin{cases} a & \text{if } (i-j) \equiv 0 \pmod 3 \\ b & \text{if } (i-j) \equiv 1 \pmod 3 \\ c & \text{if } (i-j) \equiv 2 \pmod 3 \end{cases}.$$

For each $(i,j) \in B \subset \mathbb{Z}^2$ let $e_{i,j}$ be the (i,j)th unit vector in $\mathbb{R}^B$, and define the *score map* to be the $\mathbb{Z}$-module homomorphism $\phi : \mathbb{Z}^B \longrightarrow \mathbb{Z}_2^2$ with $\phi(e_{i,j}) := (g_1(i,j)\ ,\ g_2(i,j))$. Thus, the score of a configuration $c \in \mathbb{Z}^B$ is given by

$$\phi(c) \quad = \quad \sum_{(i,j)\in B} c_{i,j} \cdot (g_1(i,j)\ ,\ g_2(i,j)) \quad .$$

Since the board B under discussion will always be clear from the context, we use the notation ϕ for any board. For instance, the score of the configuration c'_0 of one peg in the center of the English 33-board is $\phi(c'_0) = (a,a)$, as is also the score of its complement c_0; see Fig. 3. The score of the board B is defined to be $\phi(B) = \phi(\mathbb{1})$. It is easy to verify that any feasible move μ on any board B has the identity score $\phi(\mu) = (e,e)$. This gives the following proposition.

Proposition 2.1 [The rule-of-three]
A necessary condition for a pair of configuration (c,c') to be feasible is that $\phi(c'-c) = (e,e)$, namely, $c'-c \in Ker(\phi)$.

Using Proposition 2.1, we can show that, besides the configuration c'_0 given in Fig. 1, the only final configuration c' with exactly one non-zero entry $c'_{i,j} = 1$ forming a feasible pair (c_0, c') are the 4 configurations given in Fig. 2. Fig. 3 shows that $\phi(c') = (a,a) = \phi(c_0)$ if c' is one of $c'_0, c'_1 \ldots c'_4$, whereas $\phi(c') \neq (a,a)$ otherwise.

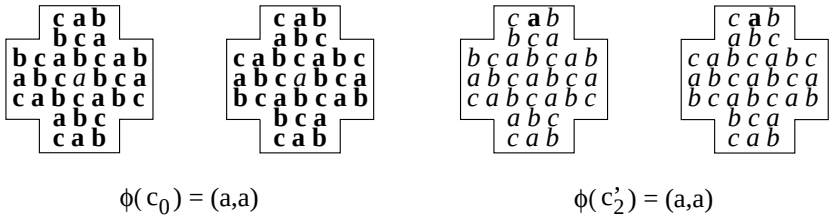

Figure 3: The score of a final configuration with only one peg remaining

With $\bar{c} = \mathbb{1} - c$ the complement of c, the rule-of-three implies that $(c, \bar{c})$ is feasible only if $\phi(\bar{c}) = \phi(c)$, which is equivalent to $\phi(B) = \phi(c) + \phi(\bar{c}) = \phi(c) + \phi(c) = (e, e)$. In other words, a necessary condition for the configurations pair $(c, \bar{c})$ to be feasible is that the board score is $\phi(B) = (e, e)$. Such a board is called a *null-class* board in [3]. For example, the score of the English 33-board is $\phi(B) = \phi(c_0) + \phi(c_0') = (a, a) + (a, a) = (e, e)$.

3 Solitaire Cone and Pagoda Functions

3.1 Solitaire cone

A first relaxation of the feasibility problem is to allow any integer (positive or negative) number of pegs to occupy any hole. We call this game the *integer game*, and call the original game the *0-1 game*. Note that in a 0-1 game we require that in every position of the game a hole is either empty or contains a single peg. We have

$$\text{c',c is } \textit{integer} \text{ feasible} \iff c' - c \in IC_B = \Big\{ \sum_{\mu \in \mathcal{M}} \lambda_\mu \mu \quad : \lambda_\mu \in \mathbb{N} \Big\}$$

where the integer cone IC_B is set of all linear combinations of moves with *non-negative integral* coefficients. Unfortunately deciding if $c' - c$ can be expressed as the sum of move seems to be a hard computational problem. A further relaxation of the game leads to a more tractable condition. In the *fractional game* we allow any hole to contain a fractional (positive or negative) number of pegs. A *fractional move* is obtained by multiplying a move by any positive scalar and is defined to correspond to the process of adding a move to a given configuration. For example, let $c =$[1 1 1], $c' =$[1 0 1]. Then $c' - c =$[0 -1 0]$=\frac{1}{2}$[-1 -1 1]$+\frac{1}{2}$[1 -1 -1] is a feasible fractional game and can be expressed as the sum of two fractional moves, but is not feasible as a 0-1 or integer game. We have:

$$\text{c',c is } \textit{fractional} \text{ feasible} \iff c' - c \in C_B = \Big\{ \sum_{\mu \in \mathcal{M}} \lambda_\mu \mu \quad : \lambda_\mu \in \mathbb{R}^+ \Big\}$$

where the *solitaire cone* C_B is set of all non-negative combinations of moves.

The condition $c' - c \in C_B$ is therefore a necessary condition for the feasibility of the original peg game and, more usefully, provides a certificate for the infeasibility of certain games. The certificate of infeasibility is any inequality valid for C_B which is violated by $c' - c$. According to [3], page 71, these inequalities "were developed by J.H. Conway and J.M. Boardman in 1961, and were called *pagoda functions* by Conway...". They are also known as *resource counts*, and are discussed in some detail in Conway [4]. The strongest such inequalities are induced by the facets of C_B. For example, the facet (iii) (given by Beasley) of Fig. 5 induces an inequality $a \cdot x \leq 0$ that is violated by $c' - c$ with (c, c') given in Fig. 4: $(c' - c) \cdot a = 2 > 0$. This implies that this game is not feasible even as a fractional game and, therefore, not feasible as an integer game or classical 0-1 game either.

3.2 Facets of the solitaire cones

Most of the results can be applied to boards which are subsets of the square lattice in the plane, such as the original Peg solitaire board. For simplicity let consider rectangular boards. For $n \geq 4$ or $m \geq 4$, the solitaire cone $C_{B_{m,n}}$ associated to the m by n board $B_{m,n}$ is a pointed

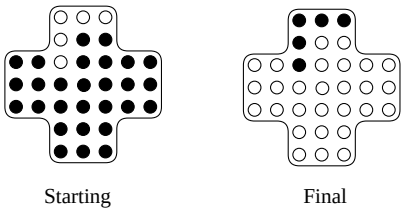

Figure 4: An infeasible classical solitaire peg game

full-dimensional cone and the moves of the solitaire cone are extreme rays; see [1] for a detailed study of $C_{B_{m,n}}$. Let assume that $m \geq 4$ or $n \geq 4$ and that the holes of $B_{m,n}$ are ordered in some way. It is convenient to display $c' - c$ and μ as m by n matrices, although of course all products should be interpreted as dot products of the corresponding mn-vectors. We represent the coefficients of the facet inducing inequality $az \leq 0$ by the m by n array $a = [a_{i,j}]$. It is a convenient abuse of terminology to refer to a as a *facet* of C_B. A *corner* of a is a coefficient $a_{i,j}$ with $i \in \{1, m\}$ and $j \in \{1, n\}$. The notation $\mathcal{T} = (t_1, t_2, t_3)$ refer to a *consecutive triple* of row or column indices. Each consecutive triple defines a *triangle inequality*: $a_{t_1} \leq a_{t_2} + a_{t_3}$ and a triangle inequality is *tight* if equality holds. The following theorem summarizes known results on properties of valid inequalities (pagoda functions) for C_B.

Theorem 3.1 [3] *For each valid inequality $a = [a_{i,j}]$ for C_B*

1. *For every consecutive triple $\mathcal{T} = (t_1, t_2, t_3)$, the associated triangle inequality must hold.*
2. *Negative coefficients of a can only occur in corners.*
3. *If $\mathcal{T} = (t_1, t_2, t_3)$ is a consecutive triple with $a_{t_2} = 0$ then $a_{t_1} = a_{t_3}$.*
4. *If two consecutive row (respectively, column) entries of a are zero the entire row (respectively, column) is zero.*

It is not feasible to generate all facets for reasonably sized boards, and in general no characterization of facets is known. A large class of facets can, however, be generated by the following procedure.

Genfacet /*procedure to generate a facet matrix a of C_B */

1. Choose a proper subset of coefficients of a satisfying: (a) If a corner is chosen, all coefficients in the row and/or column of at length at least 4 containing the corner must also be chosen; and (b) If two consecutive coefficients are chosen, their entire row and column must also be chosen. Set these chosen coefficients to zero.
2. Choose any undefined coefficient that is not a corner and set it to one.
3. Choose a consecutive triple $\mathcal{T} = (t_1, t_2, t_3)$ for which precisely two of the corresponding coefficients of a are defined. Define the remaining coefficient by the associated tight triangle inequality providing this does not violate any other triangle inequality for a.
4. Repeat step 3 until no further coefficient of a can be defined.

Theorem 3.2 [1] *Given an m by n board B, with $m \geq 4$ or $n \geq 4$, if Genfacet terminates with all elements of a defined, then a is a facet of C_B.*

Genfacet can easily be adapted to non-rectangular boards that are connected subsets of the square grid, such as the original peg solitaire game. The notion of corner generalizes in the obvious way to all holes that have exactly one horizontal and vertical neighbour. For example, the original English game has 8 corners and by Genfacet we can generate the facets (i) and (ii) given in Fig. 5:

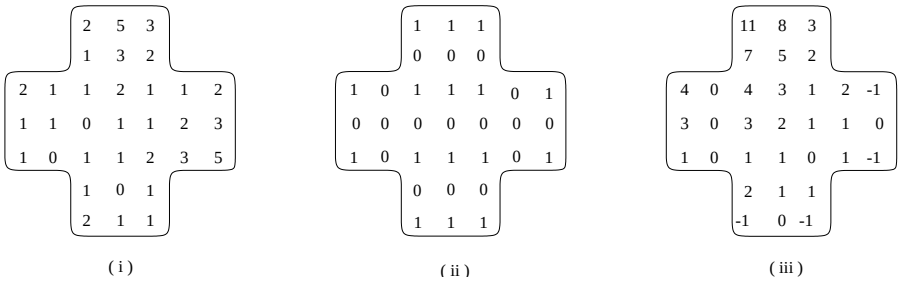

Figure 5: Three facets of the English solitaire cone

4 The Lattice Criterion

4.1 The solitaire lattice

Given a board B and an associated set of moves $\mathcal{M}$, the *solitaire lattice* of B is the lattice of all integer linear combinations of moves,

$$L_B = \Big\{ \sum_{\mu \in \mathcal{M}} \lambda_\mu \mu \quad : \lambda_\mu \in \mathbb{Z} \Big\}$$

The following proposition is immediate from the definition.

Proposition 4.1 [The lattice criterion]
A necessary condition for a pair of configurations (c, c') to be feasible is that $c' - c \in L_B$.

Since ϕ is a homomorphism of $\mathbb{Z}$-modules which maps each lattice generator $\mu \in \mathcal{M}$ to (e, e), it follows that $\phi(v) = (e, e)$ for any $v \in L_B$; i.e., for any board B and any pair (c, c') on B if $c' - c \in L_B$ then $c' - c \in \mathrm{Ker}(\phi)$. In other words, as stated in the following proposition, the lattice criterion is generally stronger than the rule-of-three.

Proposition 4.2 *For any board B, we have $L_B \subseteq Ker(\phi)$.*

Fig. 6 provides an example of a null-class board and a game on it whose associated pair $(c, \bar{c})$ satisfies $\bar{c} - c \in \mathrm{Ker}(\phi)$ but $\bar{c} - c \notin L_B$; see Section 4.3. This shows that the lattice criterion may be strictly stronger than the rule-of-three and therefore could be more useful in proving infeasibility.

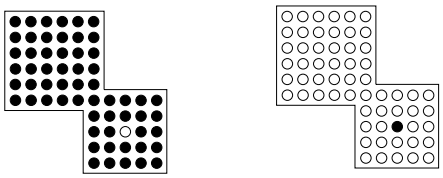

Figure 6: An infeasible game satisfying the rule-of-three but not the solitaire lattice criterion

4.2 The lattice criterion versus the rule-of-three

The solitaire lattice of a board B is typically of full rank $|B|$ and, in this case, let $\det(L_B)$ denote his determinant. Since $L_B \subseteq \mathrm{Ker}(\phi)$ for any board B, we have $\mathbb{Z}^B/\mathrm{Ker}(\phi) \subseteq \mathbb{Z}^B/L_B$ and therefore, $|\mathbb{Z}^B/\mathrm{Ker}(\phi)| \leq |\mathbb{Z}^B/L_B|$. For usual board $|\mathbb{Z}^B/\mathrm{Ker}(\phi)| = 16$: one can easily check that the 16 possible $\{0,1\}$-configurations on a 2×2 board are mapped by ϕ precisely onto the 16 elements of $\mathbb{Z}_2^2$. In other words, if a board B contains a 2×2 sub-board then $|\mathbb{Z}^B/\mathrm{Ker}(\phi)| = |\mathrm{Im}(\phi)| = 16$. For a typical board B, the map ϕ is onto and the lattice L_B is full rank, giving $|\mathbb{Z}^B/\mathrm{Ker}(\phi)| = 16$ and $|\mathbb{Z}^B/L_B| = \det(L_B)$. This gives the following useful lemma.

Lemma 4.3 *For a board B such that ϕ is an onto map and L_B is full rank, $L_B = Ker(\phi)$ if and only if $\det(L_B) = 16$.*

4.3 The lattice criterion beats the rule-of-three

A close study of $L_{B_{m,n}}$ provides a canonical basis (which is precisely the *Hermite basis*) for this lattice, see [5] for details. We have:

Theorem 4.4 *Let $B_{m,n}$ be any $m \times n$ board with $n \geq 4$ or $m \geq 4$. The solitaire lattice $L_{B_{m,n}}$ has full rank, $det(L_{B_{m,n}}) = 16$, hence $L_{B_{m,n}} = Ker(\phi)$. $L_{B_{m,n}}$ is characterized by:*

$$c \in L_{B_{m,n}} \iff \begin{cases} \sum\{c_{i,0} \ : \ 0 \leq i \leq n-1, \ i \not\equiv 0 \ (mod\ 3) \ \} & \equiv 0 \ (mod\ 2) \\ \sum\{c_{i,0} \ : \ 0 \leq i \leq n-1, \ i \not\equiv 1 \ (mod\ 3) \ \} & \equiv 0 \ (mod\ 2) \\ \sum\{c_{i,1} \ : \ 0 \leq i \leq n-1, \ i \not\equiv 0 \ (mod\ 3) \ \} & \equiv 0 \ (mod\ 2) \\ \sum\{c_{i,1} \ : \ 0 \leq i \leq n-1, \ i \not\equiv 1 \ (mod\ 3) \ \} & \equiv 0 \ (mod\ 2) \end{cases}$$

Using the basis of $L_{B_{m,n}}$, one can efficiently compute the (Hermite normal) basis of more complex boards. The algorithm consists in covering a board B by overlapping rectangular sub-boards B_i, $i = 1, \ldots, I$. Then we append each matrix $A^i_{m_i,n_i}$ $i = 1, \ldots, I$ consisting of the (row and column) moves within each sub-board B_i and add the *cross moves*, that is, the moves from one sub-board to another. Since the number of cross moves is, in general, quite small, the computation of the resulting (Hermite normal) basis of the initial board B is, in general, a bit tedious but easy. This algorithm also provides a rough upper bound for the determinant, that is, $\det(L_B) \leq \prod_{i=1}^{I} det(L_{B_i}) \leq 2^{4I}$. For example, the board B_{60} given in Fig. 6 is made of two overlapping square boards $B_{6,6}$ and $B_{5,5}$ with 4 cross-moves centered on the common hole. It directly gives the following proposition which, in particular, excludes any complementary game $(c, \bar{c})$ while, B_{60} being a null-class board, complementary games satisfy the rule-of-three.

Proposition 4.5 *L_{B60} has full rank with determinant $det(L_{B60}) = 32$, hence $L_{B60} \neq Ker(\phi)$, and $\mathbb{1} \notin L_{B60}$, that is, any complementary game $(c, \bar{c})$ is infeasible on B_{60}.*

The computation the basis of L_{B60} can be easily extended to any set of k rectangular boards pairwise overlapping on a common corner. The resulting k-hook board - see for example Fig. 7 - will give a lattice with determinant 2^{k+3}. Hook boards demonstrate that the solitaire lattice criterion can be *exponentially finer* than the rule-of-three in the sense of the following theorem.

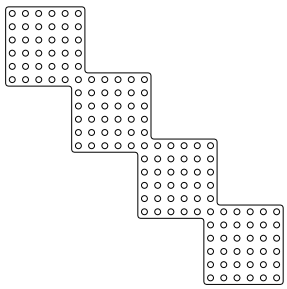

Figure 7: A hook board B with associated lattice satisfying $\det(L_B) = 128$

Theorem 4.6 *For hook boards, the solitaire lattice condition exponentially outperforms the rule-of-three, that is, for every k and every k-hook board B, the ratio of the number of congruence classes of $\mathbb{Z}^B$ modulo L_B to the number of congruence classes of $\mathbb{Z}^B$ modulo $Ker(\phi)$ satisfies*

$$\frac{[\mathbb{Z}^B / L_B]}{[\mathbb{Z}^B / Ker(\phi)]} \quad = \quad 2^{k-1} \quad .$$

References

[1] Avis D. and Deza A.: On the solitaire cone and its relationship to multi-commodity flows. *Preprint CAMS* **142** Ecole des Hautes Etudes en Sciences Sociales, Paris, France (1997)

[2] D. Avis and A. Deza: On the boolean solitaire cone. Proc. of the 1^{st} Japanese-Hungarian Symposium on Discrete Mathematics, Kyoto (1999) 115-122

[3] Beasley J. D.: The Ins and Outs of Peg Solitaire. Oxford University Press (1992)

[4] Conway J. H. with Berlekamp E. R. and Guy R. K.: Purging Pegs Properly. Winning Ways for your mathematical plays. **2** Academic Press (1982) 697-734

[5] Deza A. and Onn S.: Solitaire lattices. Research Reports on Math. and Comput. Sciences **B-345** Tokyo Institute of Technology (1998)

[6] Leibniz G. W.: Miscellenea Berolinensia ad incrementum scientarium **1** (1710) 22-26

[7] Uehara R. and Iwata S.: Generalized Hi-Q is NP-complete. Trans. IEICE **E 73** (1990) 270-273

[8] Vallot J.N.: Rapport sur un travail de Suremain de Missery: Théorie générale du jeu de solitaire, considéré comme problème d'analyse et de situation. Compte-rendu des travaux de l'académie des sciences, arts et belles-lettres de Dijon (1841-2) 58-70

David Avis. McGill University, school of computer science, 3480 University, Montréal, H3A 2A7, Canada. *Email: avis@mutt.cs.mcgill.ca*

Antoine Deza. Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo, Japan. *Email: deza@is.titech.ac.jp* and Ecole des Hautes Etudes en Sciences Sociales, Centre d'Analyse et de Mathématique Sociales, 54 bd Raspail 75270 Paris, France. *Email: deza@ehess.fr*

Shmuel Onn. Technion - Israel Institute of Technology, Operations Research, Davidson faculty of IE & M, 32000 Haifa, Israel. *Email: onn@ie.technion.ac.il*