INTEGER CONSTRAINTS FOR TRAIN SERIES CONNECTIONS

ROB A. ZUIDWIJK, LEO G. KROON

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2000-05-LIS
Publication status / version	draft / version January 2000
Number of pages	13
Email address first author	rzuidwijk@fbk.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam PoBox 1738 3000 DR Rotterdam, The Netherlands Phone: # 31-(0) 10-408 1182 Fax: # 31-(0) 10-408 9020 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	The scheduling of train services is subject to a number of constraints describing railway infrastructure, required train services and reasonable time-intervals for waiting and transits. Timetable planners at Dutch Railways are nowadays supported by a software tool, called CADANS, which produces a feasible timetable on an hourly basis. In this paper, connection requirements between train series are written in the format of the CADANS model. It turns out that this leads to nontrivial combinatorial scheduling issues.	
Library of Congress Classification (LCC)	5001-6182	Business
	5001-6182	Business Science
	331-380	Traffic engineering
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	R 4	Transportation Systems
	R 49	Transportation Systems: Railways
European Business Schools Library Group (EBSLG)	85 A	Business General
	260 K 240 B	Logistics Information Systems Management
	260 Q	Transportation
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.34	Logistiek management
	55.30	Railverkeer: algemeen
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Bedrijfsprocessen, logistiek, management informatiesystemen	
	Scheduling, Spoorwegen, Reistijden	
Free keywords	Train Series Connection, Cyclic Timetable, Railway Scheduling, MILP Constraint, Periodic Set	
Other information		

Integer Constraints for Train Series Connections

Rob A. Zuidwijk* Leo G. Kroon†

April 27, 2000

Abstract

The scheduling of train services is subject to a number of constraints describing railway infrastructure, required train services and reasonable time-intervals for waiting and transits. Timetable planners at Dutch Railways are nowadays supported by a software tool, called CADANS, which produces a feasible timetable on an hourly basis. In this paper, connection requirements between train series are written in the format of the CADANS model. It turns out that this leads to nontrivial combinatorial scheduling issues.

1 Introduction

An important step in the Dutch railway timetabling problem involves the allocation of departure and arrival times of trains on an hourly basis. Once such a feasible schedule of train services has been established, adjustments can be made in order to account for rush hours, weekends and other events. In short, the timetable for one hour (called the Basic Hourly Pattern) is a preliminary version of the final timetable which is actually used by Dutch Railways. There are several reasons for the development of the timetable in steps as mentioned above. Indeed, the sheer complexity of the problem as a whole does not invite a direct approach. The development of merely the Basic Hourly Pattern still is a difficult problem, but it is conceivable that a feasible pattern can be generated in an automated fashion for the Dutch railway system. Further adjustments of the Basic Hourly Pattern are not always well-defined and sometimes ad-hoc. Therefore, the development of the final timetable from the Basic Hourly Pattern requires intervention by a human planner.

*Rotterdam School of Management, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, E-mail: R.Zuidwijk@fbk.eur.nl

†Rotterdam School of Management, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, E-mail: L.Kroon@fbk.eur.nl

This paper deals with issues concerning the automated development of the Basic Hourly Pattern (BHP). The BHP provides a cylic timetable with period equal to one hour. Since the generation of a feasible BHP is NP-complete, it certainly is a challenge to come up with a feasible pattern of train services. Moreover, besides restrictions imposed by railway infrastructure and required line services, there are further specifications such as reasonable waiting- and transit time-intervals for passengers. Not all such specifications have yet been incorporated. In this paper, we shall deal with the problem of automating waiting- and transit time requirements for connecting train series. A train series provides a connection between two train stations (including possible intermediate stops) on a regular time basis. As we are in a BHP, such a train series is also specified by an hourly frequency. For example, a train series may provide the same train connection three times per hour.

The development of the BHP is automated by using a software tool called CADANS; see Schrijver & Steenbeek [5]. In this paper, we put attention to one aspect of the formal description of the model which underlies the CADANS-methods. References to other work in this direction are Kroon & Peeters [1], Odijk [3], Peeters [4] and Serafini & Ukovich [6].

Train services are characterized by time and place of departure and arrival. Indeed, the departure and arrival time of train t at node n is denoted by $v_{t,n}$ and $a_{t,n}$, respectively. We remark that t runs through a finite set of train numbers, and that n runs through the set of nodes in the railway network. Such a node coincides with a train station or another relevant point in the network. Since in the following argument, there is no disturbing ambiguity concerning trains or nodes involved, we shall not pursue this extensive notation.

2 Train Series Connections

We consider train series which are characterized by their time points of arrival or departure. These time points are identical from hour to hour, so the collection of time points is periodical with period $T = 60$. (The unit of time is chosen equal to one minute.) An explanation of periodic sets is given in Appendix A. Let $0 \leq a_1 < a_2 < \ldots < a_n < T$ be integers. The periodic set

$$A = \{a_1, a_2, \ldots, a_n\}_T$$

denotes the arrival times of the first train series. The arrival pattern will be evenly spread in time in e.g. the following sense:

$$a_k - a_{k-1} \in [\frac{T}{n} - \delta, \frac{T}{n} + \delta]_T, \quad k = 1, \ldots, n, \tag{1}$$

where $a_0 := a_n$. The *tolerance* δ satisfies $0 \leq 2n\delta < T$, so that the periodized intervals in (1) are mutually disjoint. The positive integer n denotes the hourly arrival frequency of the first train series. The periodic set of departure times for the second train series

$$V = \{v_1, v_2, \ldots, v_m\}_T$$

satisfies e.g.

$$v_l - v_{l-1} \in [\frac{T}{m} - \eta, \frac{T}{m} + \eta]_T, \quad l = 1, \ldots, m, \tag{2}$$

where $v_0 := v_m$ and η satisfies $0 \leq 2m\eta < T$. The positive integer m denotes the hourly departure frequency of the second train series.

In words, the conditions (1) and (2) ensure, up to some positive tolerance (defined by δ and η, respectively), that the arrival and departure times are "evenly spread in time with specified hourly frequency" (defined by n and m, respectively). For example, if one requires the arriving train series to arrive each quarter of an hour or so, then one should put $n = 4$. The conditions (1) and (2) are presently used in the CADANS software. However, there are other formal conditions that are compatible with the intuitive requirement of evenly distributing train series arrivals or departures in time. These variants are discussed in Appendix B), and are mutually compared. In particular, a more restrictive variant of condition (1) reads

$$a_k - a_l \in [\frac{(k-l)T}{n} - \delta, \frac{(k-l)T}{n} + \delta]_T, \quad k, l = 1, \ldots, n. \tag{3}$$

In the notation of Appendix B and in the following, condition (1) defines periodic sets of class $\mathcal{A}_{n,\delta}$, and (3) defines periodic sets of class $\mathcal{C}_{n,\delta}$.

The following constraint on the aforementioned arrival and departure trains defines a "connection" between train series and is of the main object of study in this paper:

(†) *There is at least one train from the first train series (with arrival time a_k) and at least one train from the second train series (with departure time v_l) such that*

$$v_l - a_k \in [\alpha, \alpha + \varepsilon]_T.$$

Here α is a positive integer indicating the minimum time required for passengers to change from the arriving train to the departing train, and the nonnegative integer ε indicates the extent up to which the transfer time can be stretched.

It is completely straightforward to make A and V satisfy this constraint; simply take $v_1 = a_1 + \alpha$. However, the underlying situation is more complicated. The sets of arrival and departure times are decision variables in a mixed integer program with a large and intricate set of constraints. It is the aim of this paper to reformulate the train series connection constraint (†) in terms of mutual differences, which is the appropriate form for the CADANS software; see Schrijver & Steenbeek [5]. The following theorem provides such a reformulation.

Theorem 2.1 *Let $A = \{a_1, \dots, a_n\}$ be of class $\mathcal{C}_{n,\delta}$ and $V = \{v_1, \dots, v_m\}$ be of class $\mathcal{C}_{m,\eta}$. Let $\alpha, \varepsilon \in [0, T)$ and assume that*

$$\varepsilon + 3\delta + 3\eta < \frac{T}{\operatorname{lcm}(m, n)}. \tag{4}$$

Write $d = \gcd(m, n)$ and $m' = m/d, n' = n/d$ and $T' = T/d$. Then the following two statements are equivalent:

1. *For $j = 1, \dots, d$, there exist $(j-1)n' \le k_j < jn'$ and $(j-1)m' \le l_j < jm'$, such that*
$$v_{l_j} - a_{k_j} \in [\alpha, \alpha + \varepsilon]_T.$$

2. *For all $0 \le k < n$ and $0 \le l < m$ such that $(j-1)n' \le k < jn'$ and $(j-1)m' \le l < jm'$ for some $j \in \{1, \dots, d\}$, it holds true that*
$$v_l - a_k \in \left(\bigcup_{\mathcal{P}} \frac{pT'}{n'} + \frac{qT'}{m'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_{T'} \right) \cup [\alpha, \alpha + \varepsilon]_T,$$
where $\mathcal{P} = \{(p, q) \mid 0 \le p < n',\ 0 \le q < m',\ (p, q) \ne (0, 0)\}$.

Proof In order to prove that the first statement implies the second one, let $0 \le k < n$ and $0 \le l < m$ such that there exists $j \in \{1, \dots, d\}$ with $k \in I_j(n') := [(j-1)n', jn')$ and $l \in I_j(m')$. It holds true that

$$v_l - a_k = (v_l - v_{l_j}) + (a_{k_j} - a_k) + (v_{l_j} - a_{k_j}) \in$$
$$\frac{(l - l_j)T'}{m'} + \frac{(k_j - k)T'}{n'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_T \subseteq$$
$$\frac{(l - l_j)T'}{m'} + \frac{(k_j - k)T'}{n'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_{T'}.$$

In the case when $l = l_j$ and $k = k_j$, we have $v_l - a_k \in [\alpha, \alpha + \varepsilon]_T$. In the case when $(k - k_j, l - l_j) \ne (0, 0)$, write $p = k_j - k + \mu n'$ for such integer μ that $0 \le p < n'$, and, in the same fashion, define $q = l - l_j + \nu m' \in [0, m')$.

Conversely, observe that there exist $dm'n'$ pairs (a_k, v_l), such that $(k, l) \in I_j(n') \times I_j(m')$ for some $j \in \{1, \dots, d\}$. Moreover, there are $m'n' - 1$ sets of the form

$$\frac{pT'}{n'} + \frac{qT'}{m'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_{T'},$$

where $(p, q) \in \mathcal{P}$.

Claim: For each $1 \le j \le d$ and $(p, q) \in \mathcal{P}$, there exists at most one pair (a_k, v_l), such that $k \in I_j(n'), l \in I_j(m')$, and

$$v_l - a_k \in \frac{pT'}{n'} + \frac{qT'}{m'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_{T'}.$$

Proof claim: Assume that $(k,l) \neq (r,s)$, $k,r \in I_j(n')$ and $l,s \in I_j(m')$, and both $v_l - a_k$ and $v_s - a_r$ are in the set

$$\frac{pT'}{n'} + \frac{qT'}{m'} + [-\delta - \eta + \alpha, \delta + \eta + \alpha + \varepsilon]_{T'}.$$

Then

$$(v_l - a_k) - (v_s - a_r) = (v_l - v_s) + (a_r - a_k) \in \frac{(l-s)T}{m} + \frac{(r-k)T}{n} + [-\delta - \eta, \delta + \eta]_T \subseteq$$

$$\frac{(l-s)T'}{m'} + \frac{(r-k)T'}{n'} + [-\delta - \eta, \delta + \eta]_{T'}.$$

On the other hand,

$$(v_l - a_k) - (v_s - a_r) \in [-2\delta - 2\eta - \varepsilon, 2\delta + 2\eta + \varepsilon]_{T'}.$$

Observe that $\frac{(l-s)T'}{m'} + \frac{(r-k)T'}{n'} = \mu T'$ with $\mu \in \mathbb{Z}$ if and only if there exists $\lambda \in \mathbb{Z}$ such that $k - r = \lambda n'$ and $l - s = (\mu - \lambda)m'$. This is in contradiction with $(k,l) \neq (r,s)$ and $|k - r| < n'$ and $|l - s| < m'$.

Note that, as a consequence, $\frac{(l-s)T'}{m'} + \frac{(r-k)T'}{n'}$ is a multiple of $\frac{T'}{m'n'}$ and not a multiple of T'. Therefore, the sets

$$\frac{(l-s)T'}{m'} + \frac{(r-k)T'}{n'} + [-\delta - \eta, \delta + \eta]_{T'}$$

and

$$[-2\delta - 2\eta - \varepsilon, 2\delta + 2\eta + \varepsilon]_{T'}$$

have nonzero intersection, only if

$$\frac{T'}{m'n'} = \frac{T}{\operatorname{lcm}(m,n)} \leq 3\delta + 3\eta + \varepsilon.$$

This is in contradiction with the assumptions in the theorem, so the claim is proved.

Using the same type of argument as in the proof of the claim, it can be shown that at most one integer pair $(k,l) \in I_j(n') \times I_j(m')$ provides a difference $v_l - a_k \in [\alpha, \alpha + \varepsilon]_T$.We conclude that for each $1 \leq j \leq d$, the set $[\alpha, \alpha + \varepsilon]_T$ contains at least $m'n' - (m'n' - 1) = 1$, and hence exactly 1, difference $v_l - a_k$ with $(k,l) \in I_j(m') \times I_j(n')$. This proves the lemma. □

The inequality (4) is close to being necessary for the statement of the theorem, as the following proposition shows. Two integers are called co-prime if their greatest common divisor is equal to one.

Proposition 2.2 *We assume here that m and n are co-prime, and that*

$$0 \le \delta, \quad 0 \le \varepsilon \le \eta, \quad \delta + 2\eta + \varepsilon \le \frac{T}{mn}, \quad 2\delta + 2\eta + \varepsilon \ge \frac{T}{mn}.$$

The sets

$$A = \left\{ \frac{kT}{n} - \delta - \eta - \varepsilon, \frac{(mn-1)T}{mn} + \delta + \eta \mid 1 \le k \le n-1 \right\}_T,$$

and

$$V = \left\{ \alpha, \alpha + \frac{lT}{m} - \varepsilon \mid 1 \le l \le m-1 \right\}_T$$

satisfy $A \in \mathcal{C}_{n,\delta}$, $V \in \mathcal{C}_{m,\eta}$. Moreover,

$$V - A \subseteq \bigcup_{k=1}^{mn-1} \alpha + \frac{kT}{mn} + [-\delta - \eta, \delta + \eta + \varepsilon]_T. \tag{5}$$

This implies that the equivalence as stated in Theorem 2.1 does not hold here.

For the proof of the proposition, we need the following rather general algebraic statement. The proof involves elementary algebra and is omitted here.

Lemma 2.3 *Let m and n be co-prime integers. Then*

$$\left\{ \frac{pT}{n} - \frac{qT}{m} \mid 1 \le p \le n,\ 1 \le q \le m \right\}_T = \left\{ \frac{kT}{mn} \mid 0 \le k \le mn-1 \right\}_T.$$

Proof of Proposition 2.2 We first prove that $A \in \mathcal{C}_{n,\delta}$. Write $a_k = \frac{kT}{n} - \delta - \eta - \varepsilon$ for $1 \le k \le n-1$ and $a_n = \frac{(mn-1)T}{mn} + \delta + \eta$. Then $a_l - a_k = \frac{(l-k)T}{n}$ for $1 \le k \le l < n$. Moreover, $a_n - a_k = \frac{(mn-1)T}{mn} - \frac{kT}{n} + 2\delta + 2\eta + \varepsilon = \frac{(n-k)T}{n} + 2\delta + 2\eta + \varepsilon - \frac{T}{mn}$. Observe that $0 \le 2\delta + 2\eta + \varepsilon - \frac{T}{mn} \le \delta$. The proof that $V \in \mathcal{C}_{m,\eta}$ is straightforward and omitted here. Next, we prove that $V - A$ is contained in the right hand side of (5). Note that

$$V - A = \left\{ \alpha - \frac{kT}{n} + \delta + \eta + \varepsilon, \alpha + \frac{T}{mn} - \delta - \eta, \frac{lT}{m} - \frac{kT}{n} + \delta + \eta, \right.$$
$$\left. \alpha - \frac{T}{mn} + \frac{lT}{m} - \delta - \eta - \varepsilon \mid 1 \le l \le n-1,\ 1 \le l \le m-1 \right\}_T,$$

while $\frac{T}{mn} + \frac{lT}{m} - \delta - \eta - \varepsilon \in \frac{lT}{m} + [\eta, \delta + \eta]_T$. Remains to prove that the inclusion (5) implies that the equivalence in Theorem 2.1 does not hold. First of all, note that

$$\mathcal{U} = \bigcup_{k=1}^{mn-1} \frac{kT}{mn} + [-\delta - \eta, \delta + \eta + \varepsilon]_T$$

and $[0,\varepsilon]_T$ have empty intersection, whenever $\delta+\eta+\varepsilon < \frac{T}{mn}$. Indeed,

$$\mathcal{U} \subseteq \left(\frac{T}{mn} - \delta - \eta, T - \frac{T}{mn} + \delta + \eta + \varepsilon\right)_T \subseteq (\varepsilon, T)_T.$$

The same holds for the translated sets $\alpha + \mathcal{U}$ and $[\alpha, \alpha+\varepsilon]_T$. Further, Lemma 2.3 implies that

$$\mathcal{U} = \bigcup_{\mathcal{P}} \frac{pT}{n} - \frac{qT}{m} + [-\delta-\eta, \delta+\eta+\varepsilon]_T,$$

with $\mathcal{P} = \{(p,q) \mid 0 \le p \le n-1,\ 0 \le q \le m-1,\ (p,q) \ne (0,0)\}$. Therefore, $V - A \subseteq \alpha + \mathcal{U}$ is equivalent to the statement that, in Theorem 2.1, statement 2) holds but statement 1) does not hold. □

Observe that the case when $T = 60$, $m = 2$, $n = 3$, $\varepsilon = \delta = \eta = 2$ (and $\alpha = 3$) is covered by Proposition 2.2. This example is one of the first interesting cases arising in the study of train series connections; see [2]. Therefore, we cannot expect a statement as in Theorem 2.1 for all relevant cases.

The preceding results concerned the reformulation of restrictions which describe train series connections. In short these restrictions impose that, between certain departure times v_l and arrival times a_k, there should be a restricted (periodic) time interval, i.e., for specific integers k, l,

$$v_l - a_k \in [\alpha, \alpha+\varepsilon]_T.$$

These restrictions do not provide a preference among feasible time differences $v_l - a_k$. In practice, however, such preferences do exist. We shall model these preferences by introducing a penalty function φ on the interval $[\alpha, \alpha+\varepsilon]$, given by $\varphi(t) = (t-\alpha)/\varepsilon$. In words, the transit time α is considered to be optimal, a shorter transfer time is not feasible, while a longer transfer time is penalized with a finite penalty up to the extent of $\alpha+\varepsilon$. Transfer times beyond this point are not considered to define a connection. The following proposition indicates that such a penalty policy can be realized using a penalty function. The penalty function is linear only in the arrival and departure variables which are involved in actual connections between the two train series. As there are only $d = \gcd(m,n)$ of such train connections, the penalty function φ is linear in all arrival and departure times, if $m = n = d$. In all other cases, the analysis of the penalty function seems more intricate and is is the object of current research.

In the following proposition, notation is consistent with the remainder of this section. We define the function $\varphi : \mathbb{R} \to [0,1]$ by

$$\varphi(t) = \begin{cases} \frac{(t \bmod T)-\alpha}{\varepsilon}, & t \in [\alpha, \alpha+\varepsilon]_T \\ 1, & \text{otherwise} \end{cases}$$

Proposition 2.4 *If the series of arrivals $\{a_k\}_{k=1}^n$ and departures $\{v_l\}_{l=1}^m$ satisfy the conditions and statements mentioned in Theorem 2.1, then*

$$\sum_{j=1}^{d} \sum_{(k,l)\in I_j(n')\times I_j(m')} \varphi(v_l - a_k) = \frac{1}{\varepsilon}\sum_{j=1}^{d} v_{l_j} - \frac{1}{\varepsilon}\sum_{j=1}^{d} a_{k_j} - \frac{\alpha d}{\varepsilon} + \mathrm{lcm}(m,n) - d,$$

where $\{a_{k_j}\}_{j=1}^d$ and $\{v_{l_j}\}_{j=1}^d$ are the arrivals and departures involved in the d connections, respectively, as described in Theorem 2.1.

Proof The first statement in theorem 2.1 states that, for each $j \in \{1,\ldots,d\}$ fixed, there exist $k_j \in [(j-1)n', jn')$ and $l_j \in [(j-1)m', jm')$, such that

$$v_{l_j} - a_{k_j} \in [\alpha, \alpha+\varepsilon]_T.$$

For arbitrary $(k,l) \in I_j(n') \times I_j(m')$, we obtain

$$v_l - a_k = (v_l - v_{l_j}) + (a_{k_j} - a_k) + (v_{l_j} - a_{k_j}) \in$$

$$\alpha + \frac{(l-l_j)T'}{m'n'} - \frac{(k-k_j)T'}{m'n'} + [\alpha-\delta-\eta, \alpha+\delta+\eta+\varepsilon]_{T'}$$

By assumption, these $m'n'$ intervals are mutually disjoint, and $[\alpha, \alpha+\varepsilon]_T$ has non-empty intersection only with $[-\delta-\eta, \delta+\eta+\varepsilon]_{T'}$. The proof of theorem 2.1 reveals that each one of the $m'n'$ intervals contains exactly one of the $m'n'$ differences $v_l - a_k$. Therefore,

$$\sum_{(k,l)\in I_j(n')\times I_j(m')} \varphi(v_l - a_k) = \frac{v_{l_j} - a_{k_j} - \alpha}{\varepsilon} + m'n' - 1.$$

The proposition has been proved. □

Observe that Proposition 2.4 rewrites the penalty function as a linear function. However, the prescription of the linear function uses knowledge of the arrivals and departures that are actually involved in the train series connections. Therefore, the proposition does in general not provide a linear penalty function to be used for all instances.

The only case when the proposition does provide a linear penalty function is the case when $m = n = d$. Indeed, the result of the proposition comes down to

$$\sum_{j=1}^{d} \varphi(v_{l_j} - a_{k_j}) = \frac{1}{\varepsilon}\sum_{l=1}^{d} v_l - \frac{1}{\varepsilon}\sum_{k=1}^{d} a_k - \frac{\alpha d}{\varepsilon}.$$

A Periodic sets

Given a period $T > 0$, we want to define sets which are periodic in the sense that they are invariant under shifts by (integer multiples of) T. A somewhat formal approach (inspired by [6]) is chosen here in order to avoid cumbersome definitions with limited generality. The formal definition will be specified to a special case of periodic intervals which is relevant for the applications at hand.

Two real numbers x and y are said to be equivalent ($x \sim y$) if there exists $n \in \mathbb{Z}$ such that $x - y = nT$. Further, we define equivalence classes $[x]$ for $x \in \mathbb{R}$ which satisfy: $[x] = [y]$ if and only if $x \sim y$. The collection of equivalence classes is denoted by $\mathbb{R}_{/\sim}$. The canonical mapping $\pi : \mathbb{R} \to \mathbb{R}_{/\sim}$ is simply given by $\pi(x) = [x]$. A converse mapping $\pi^+ : \mathbb{R}_{/\sim} \to \mathcal{P}(\mathbb{R})$ is defined by $\pi^+([x]) = \{y \in \mathbb{R} \mid y \sim x\}$. Here $\mathcal{P}(\mathbb{R})$ denotes the collection of all subsets in $\mathbb{R}$. We make the following simple observation.

Lemma A.1 *Let π and π^+ be defined as above. Then*

$$\pi(\pi^+([x])) = [x] \quad \text{for all} \quad [x] \in \mathbb{R}_{/\sim},$$

$$\pi^+(\pi(x)) = \{y \in \mathbb{R} \mid y \sim x\} \quad \text{for all} \quad x \in \mathbb{R}.$$

Given $A \subseteq \mathbb{R}$, we introduce its *periodization*

$$A_T = \pi^+(\pi(A)).$$

A subset A is called *periodic* if $A = A_T$. As a simple example, let $[a, b] \subseteq \mathbb{R}$ be a compact interval. By definition, we get

$$[a, b]_T = \{x \in \mathbb{R} \mid \text{there exists } n \in \mathbb{Z} \text{ such that } a \leq x + nT \leq b\}.$$

Observe that if $n \in \mathbb{Z}$, then $[a, b]_T = [a + nT, b + nT]_T$. Therefore, without loss of generality, we may assume that $0 \leq a < T$.

In the paper, the following lemma is used frequently. The sum of two sets is defined as

$$A + B = \{a + b \mid a \in A,\ b \in B\}.$$

Lemma A.2 *Let $a_1 \leq b_1$ and $a_2 \leq b_2$ be real numbers, then*

$$[a_1, b_1]_T + [a_2, b_2]_T = [a_1 + a_2, b_1 + b_2]_T.$$

Proof We first prove that the inclusion"$\subseteq$" holds. If $x_j \in [a_j, b_j]_T$, then there exists $n_j \in \mathbb{Z}$ such that $a_j \leq x_j - n_jT \leq b_j$ ($j = 1, 2$). Therefore, $a_1 + a_2 \leq x_1 + x_2 - (n_1 + n_2)T \leq b_1 + b_2$. The other inclusion is proved as follows. Assume that $x \in [a_1 + a_2, b_1 + b_2]_T$, i.e., there exists $n \in \mathbb{Z}$ such that $z = x - nT \in [a_1 + a_2, b_1 + b_2]$. We need to find $y \in \mathbb{R}$ such that $y \in [a_1, b_1]$ and $z - y \in [a_2, b_2]$. Define the rectangle $[a_1, b_1] \times [a_2, b_2]$ in the real plane. Observe that this rectangle intersects the line $\{(u, v) \mid u + v = z\}$. Define $y \in \mathbb{R}$ in such a way that the point $(y, z - y)$ is in this intersection. □

We have introduced the notions of periodization and periodicity over the real numbers. In the paper, we shall use these notions over the integers only. Accordingly, the period T will be a strictly positive integer.

B Classes of constraints

We consider three classes of periodic sets. Let $A = \{a_1, \ldots, a_n\}$ consist of n distinct integers in $[0, T)$. Fix a (small) positive integer δ such that $2n\delta < T$. We assume that n is a divisor of T.

1. A is from the class $\mathcal{A}_{n,\delta}$, if $(a_0 := a_n)$

$$a_k - a_{k-1} \in \frac{T}{n} + [-\delta, \delta]_T, \quad k = 1, \ldots, n.$$

2. A is from the class $\mathcal{B}_{n,\delta}$, if there exists $\tau \in [0, T)$, such that

$$a_k \in \tau + \frac{kT}{n} + [-\delta, \delta]_T, \quad k = 1, \ldots, n.$$

3. A is from the class $\mathcal{C}_{n,\delta}$, if

$$a_k - a_l \in \frac{(k - l)T}{n} + [-\delta, \delta]_T, \quad k, l = 1, \ldots, n.$$

The three classes are related in a somewhat intricate fashion. The following lemmas provide a detailed description of these matters.

Lemma B.1 *The following statements hold:*

$$\mathcal{C}_{n,\delta} \subseteq \mathcal{A}_{n,\delta}, \quad n \geq 1,$$
$$\mathcal{C}_{n,\delta} = \mathcal{A}_{n,\delta}, \quad 1 \leq n \leq 3,$$
$$\mathcal{C}_{n,\delta} \neq \mathcal{A}_{n,\delta}, \quad n \geq 4.$$

Proof The inclusion $\mathcal{C}_{n,\delta} \subseteq \mathcal{A}_{n,\delta}$ for $n \geq 1$ follows immediately from a comparison of the definitions of $\mathcal{A}_{n,\delta}$ and $\mathcal{C}_{n,\delta}$. The equality for $1 \leq n \leq 3$ is straightforward. In order to prove that $\mathcal{C}_{n,\delta} \neq \mathcal{A}_{n,\delta}$ for $n \geq 4$, we define

$$A = \left\{0, \frac{T}{n} + \delta, \frac{2T}{n} + 2\delta, \frac{3T}{n} + \delta, \frac{4T}{n}, \ldots, \frac{(n-1)T}{n}\right\}_T.$$

Indeed, it is easy to see that $A \in \mathcal{A}_{n,\delta}$ while $A \notin \mathcal{C}_{n,\delta}$. □

Lemma B.2 *The following statements hold:*

$$\mathcal{C}_{n,\delta} \subseteq \mathcal{B}_{n,\delta}, \quad n \geq 1,$$

$$\mathcal{C}_{n,\delta} \neq \mathcal{B}_{n,\delta}, \quad n \geq 2.$$

Proof Let $A = \{a_1, \ldots, a_n\}$ be in the class $\mathcal{C}_{n,\delta}$ and put $\tau = a_n$. Then $a_n \in \tau + [-\delta, \delta]_T$ and if $k \neq n$, we get

$$a_k - \tau = a_k - a_n \in \frac{kT}{n} + [-\delta, \delta]_T,$$

so $a_k \in \tau + \frac{kT}{n} + [-\delta, \delta]_T$. This proves $A \in \mathcal{B}_{n,\delta}$. On the other hand, if $n \geq 2$, the set

$$A = \{\delta, \frac{T}{n} - \delta, \frac{2T}{n}, \ldots, \frac{(n-1)T}{n}\}$$

is in $\mathcal{B}_{n,\delta}$, but not in $\mathcal{C}_{n,\delta}$. □

Lemma B.3 *The following statements hold:*

$$\mathcal{C}_{n,\delta} \subseteq \mathcal{A}_{n,\delta} \cap \mathcal{B}_{n,\delta}, \quad n \geq 1,$$

$$\mathcal{C}_{n,\delta} = \mathcal{A}_{n,\delta} \cap \mathcal{B}_{n,\delta}, \quad 1 \leq n \leq 3,$$

$$\mathcal{C}_{n,\delta} \neq \mathcal{A}_{n,\delta} \cap \mathcal{B}_{n,\delta}, \quad n \geq 4.$$

Proof The first statement follows immediately from Lemmas B.1 and B.2. The second statement follows from the fact that $\mathcal{C}_{n,\delta} = \mathcal{A}_{n,\delta}$ for $n \in \{2, 3\}$ and from the first statement. The third statement is proved using the example

$$A = \left\{0, \frac{T}{n} - \delta, \frac{2T}{n}, \frac{3T}{n} + \delta, \frac{4T}{n}, \ldots, \frac{(n-1)T}{n}\right\}_T.$$

Indeed, $A \in \mathcal{A}_{n,\delta} \cap \mathcal{B}_{n,\delta}$, but $A \notin \mathcal{C}_{n,\delta}$. □

Lemma B.4 *If $n \geq 2$, then*

$$\mathcal{B}_{n,\delta} \not\subseteq \mathcal{A}_{n,\delta}. \tag{6}$$

For $2 \leq n \leq 5$, we get

$$\mathcal{A}_{n,\delta} \subset \mathcal{B}_{n,\delta}, \tag{7}$$

while for $n \geq 6$, we have

$$\mathcal{A}_{n,\delta} \not\subseteq \mathcal{B}_{n,\delta}. \tag{8}$$

Proof For $n \geq 2$, the set (see Lemma B.2)

$$\{\delta, \frac{T}{n} - \delta, \frac{2T}{n}, \ldots, \frac{(n-1)T}{n}\}$$

is in $\mathcal{B}_{n,\delta}$ (take $\tau = 0$), but not in $\mathcal{A}_{n,\delta}$. This proves (6). We verify the inclusion (7) as follows: Let $A = \{a_1, \ldots, a_n\} \in \mathcal{A}_{n,\delta}$, and write

$$\varepsilon_k = a_k - \frac{kT}{n} - \nu_k T, \nu_k \in \mathbb{Z},$$

in such a way that $0 \leq \varepsilon_k < T$ for $k = 1, \ldots, n$. By definition ($\varepsilon_0 := \varepsilon_n$),

$$\varepsilon_k - \varepsilon_{k-1} \in [-\delta, \delta]_T, \quad k = 1, \ldots, n.$$

Define $\varepsilon_+ = \max\{\varepsilon_k \mid 1 \leq k \leq n\}$ and $\varepsilon_- = \min\{\varepsilon_k \mid 1 \leq k \leq n\}$. Since $n \leq 5$, it holds that $\varepsilon_+ - \varepsilon_- \in [-2\delta, 2\delta]_T$ and henceforth $0 \leq \varepsilon_+ - \varepsilon_- \leq 2\delta$. Put $\tau = \frac{\varepsilon_+ + \varepsilon_-}{2}$, then

$$|\tau - \varepsilon_k| \leq \frac{|\varepsilon_+ - \varepsilon_k| + |\varepsilon_k - \varepsilon_-|}{2} \leq \delta,$$

since $|\varepsilon_+ - \varepsilon_k| + |\varepsilon_k - \varepsilon_-| = \varepsilon_+ - \varepsilon_-$. For $n \geq 6$, a different situation arises: The set

$$\{0, \frac{T}{n} - \delta, \frac{2T}{n} - 2\delta, \frac{3T}{n} - \delta, \frac{4T}{n}, \frac{5T}{n} + \delta, \frac{6T}{n}, \ldots, \frac{(n-1)T}{n}\}$$

is in $\mathcal{A}_{n,\delta}$ but not in $\mathcal{B}_{n,\delta}$. □

Lemma B.5 *For $n \geq 1$, we get*

$$\mathcal{B}_{n,\delta} \subseteq \mathcal{C}_{n,2\delta}.$$

Proof Assume that $B = \{b_1, \ldots, b_n\} \in \mathcal{B}_{n,\delta}$, then

$$b_l - b_k \in \tau + \frac{kT}{n} - \tau - \frac{lT}{n} + [-2\delta, 2\delta]_T, \quad 1 \leq k, l \leq n.$$

□

Lemma B.6 $\mathcal{A}_{n,\delta} \subseteq \mathcal{C}_{n,\eta}$ *if and only if* $\eta \geq \lfloor \frac{n}{2} \rfloor \delta$.

Proof Assume that $\kappa = \lfloor \frac{n}{2} \rfloor$ satisfies $\kappa\delta > \eta$. Define the set A, for n odd, by

$$A = \{0, \delta + \frac{T}{n}, 2\delta + \frac{2T}{n}, \ldots, \kappa\delta + \frac{\kappa T}{n}, (\kappa-1)\delta + \frac{(\kappa+1)T}{n}, \ldots, \delta + \frac{(2\kappa-1)T}{n}, \frac{2\kappa T}{n}\},$$

and, for n even, by

$$A = \{0, \delta + \frac{T}{n}, 2\delta + \frac{2T}{n}, \ldots, \kappa\delta + \frac{\kappa T}{n}, (\kappa-1)\delta + \frac{(\kappa+1)T}{n}, \ldots, \delta + \frac{(2\kappa-1)T}{n}.\}$$

Then for all $n \geq 1$, $A \notin \mathcal{C}_{n,\eta}$, while $A \in \mathcal{A}_{n,\delta}$.

On the other hand, if $\eta \geq \lfloor \frac{n}{2} \rfloor \delta = \kappa\delta$, and $A = \{a_1, \ldots, a_n\} \in \mathcal{A}_{n,\delta}$, we find that for $1 \leq k, l \leq n$,

$$a_l - a_k \in \frac{(l-k)T}{n} + [-|l-k|\delta, |l-k|\delta]_T.$$

Without loss of generality, we may assume that $k \leq l$. Moreover, either $l - k \leq \lfloor \frac{n}{2} \rfloor$ or $k + n - l \leq \lfloor \frac{n}{2} \rfloor$. In both cases,

$$a_l - a_k \in \frac{(l-k)T}{n} + [-\kappa\delta, \kappa\delta]_T \subseteq \frac{(l-k)T}{n} + [-\eta, \eta]_T.$$

□

References

[1] L.G. Kroon, L.W.P. Peeters, A variable running times model cyclic railway timetabling, Management Report 28-1999, Rotterdam School of Management, 1999.

[2] H. v.d. Lek, P.W.F. Alons, Treinseries, februari 1996 (in Dutch and unpublished).

[3] M.A. Odijk, Railway Timetable Generation, PhD thesis, Delft University of Technology, January 1998.

[4] L.W.P. Peeters, An optimization approach to railway timetabling, contribution to: Fourth Workshop on Models and Algorithms for Planning and Scheduling Problems, Renesse, The Netherlands, 1999.

[5] A. Schrijver, A. Steenbeek, Dienstregelingontwikkeling voor Railned, Rapport CADANS 1.0, Centrum voor wiskunde en Informatica, december 1994.

[6] P. Serafini, W. Ukovich, A Mathematical Model for Periodic Scheduling Problems, *SIAM J. Discrete Mathematics* 2(4): 550-581, November 1989.

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

Other Publications in the Report Series Research* in Management

Impact of the Employee Communication and Perceived External Prestige on Organizational Identification
Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn
ERS-2000-01-MKT

Critical Complexities, from marginal paradigms to learning networks
Slawomir Magala
ERS-2000-02-ORG

Forecasting Market Shares from Models for Sales
Dennis Fok & Philip Hans Franses
ERS-2000-03-MKT

A Greedy Heuristic for a Three-Level Multi-Period Single-Sourcing Problem
H. Edwin Romeijn & Dolores Romero Morales
ERS-2000-04-LIS

---

* ERIM Research Programs:
- LIS Business Processes, Logistics and Information Systems
- ORG Managing Relationships for Performance
- MKT Decision Making in Marketing Management
- F&A Financial Decision Making and Accounting
- STR Strategic Renewal and the Dynamics of Firms, Networks and Industries