A Dynamic Approach to Timed Behavior

Jeremy Gunawardena
External Research Program
HPL-94-41
April, 1994

jhcg@hplb.hpl.hp.com

cycle time, dynamical system, max-plus algebra, min-max function, real-time reactive system

In classical applied mathematics time is treated as an independent variable. The equations which govern the behavior of a system enable one to determine, in principle, what happens at a given time. This approach has led to powerful techniques for calculating numerical quantities of interest in engineering. The temporal behavior of reactive systems, however, is difficult to treat from this viewpoint: time is usually a dependent variable—packets are timestamped in a distributed system—and it has been customary to rely on logical rather than dynamic methods. In this paper we describe the theory of min-max functions which permits a dynamic approach to the time behavior of a restricted class of reactive systems. These systems arise in practice in the timing analysis of asynchronous circuits, where, in the absence of a clock, it is important to find some measure of the speed of a circuit. The cycle time vector of a min-max function provides the appropriate measure. In this paper we study a fundamental question, the Duality Conjecture, whose affirmative answer would yield a formula for calculating the cycle time vector. Our main result is a proof of the Conjecture for min-max functions of dimension 2. The work described here was done as part of project STETSON, a joint project between HP Labs and Stanford University on asynchronous hardware design.

1 Introduction

In classical applied mathematics time is treated as an independent variable. The equations which govern the behaviour of a system enable one to determine, in principle, what happens at a given time. A convenient abstract formulation of this approach is the idea of a dynamical system, (S, F), where S is a set and $F : S \to S$ is a self-map of S, [5]. The basic problem in dynamical systems[1] is to understand the behaviour of the sequence

$$x,\ F(x),\ F^2(x),\ \cdots \tag{1}$$

for different $x \in S$. If the elements of S are thought of as states of the system then $F^s(x)$ tells us the state at time s if the system is started in state x at time 0. Although it is now well understood that an analytic (closed form) solution for $F^s(x)$ is only possible in very isolated special cases, there has been much progress in understanding the asymptotic behaviour of the system. That is, in understanding what happens when $s \to \infty$. The dynamic approach has led to powerful techniques for calculating quantities of interest in applications.

The temporal behaviour of reactive systems, however, is difficult to treat from a dynamic viewpoint: time is usually a dependent variable—packets are timestamped in a distributed system—and it has been customary to rely on logical rather than dynamic methods. This has resulted in the development of several theories of timed behaviour: various species of timed process algebra, timed I/O automata, real time temporal logics and timed Petri nets. (An overview of these approaches may be found in [17].) They have been used to specify and to reason about the behaviour of real time systems but, for the most part, it has not been easy to use them to calculate the numbers that an engineer would like to know: throughput, latency, cycle time, minimum buffer capacity, etc.

In this paper we shall show that for a restricted class of reactive systems it is possible to use dynamic methods and that numerical quantities of interest in applications can be calculated. Our approach is built on the theory of min-max functions which was first introduced in [10] following earlier work in [15]. (For a discussion of the history of this area, see [7, §1].) The following two definitions introduce the basic idea.

Let $a \vee b$ and $a \wedge b$ denote the maximum (least upper bound) and minimum (greatest lower bound) of real numbers respectively: $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. It is well known that these operations are associative and commutative and that each distributes over the other. Furthermore, addition distributes over both maximum and minimum:

$$h + (a \vee b) = h + a \vee h + b, \qquad h + (a \wedge b) = h + a \wedge h + b. \tag{2}$$

In expressions such as these $+$ has higher binding than $\wedge$ or $\vee$.

Definition 1.1 *A min-max expression, f, is a term in the grammar:*

$$f := x_1, x_2, \cdots \mid f + a \mid f \wedge f \mid f \vee f$$

where $x_1, x_2, \cdots$ are variables and $a \in \mathbf{R}$.

For example, $x_1 + x_2 \wedge x_3 + 2$ and $x_1 \vee 2$ are forbidden but $x_1 - 1 \vee x_2 + 1$ is allowed. The numbers which appear in a min-max expression are called parameters.

[1]This is a discrete dynamical system. Continuous dynamical systems derived from differential equations can be made discrete by considering the time 1 map of the corresponding flow or by the method of Poincaré sections. In this paper we need only consider discrete systems.

Definition 1.2 *A min-max function of dimension n is any function, $F : \mathbf{R}^n \to \mathbf{R}^n$, each of whose components, $F_i : \mathbf{R}^n \to \mathbf{R}$, is a min-max expression of n variables $x_1, \cdots, x_n$.*

As described above, the main problem in the theory of min-max functions is to study the asymptotic behaviour of the sequence (1) where S is Euclidean space of n dimensions and F is a min-max function of dimension n. To understand what needs to be calculated, it is helpful to consider the applications in which min-max functions arise. They were introduced to deal with the problem of defining the speed of an asynchronous digital circuit, [11]. In the absence of a clocking mechanism, a design engineer needs some measure by which circuits performing the same function can be compared in speed of operation. The behaviour of certain asynchronous circuits can be represented by a min-max function in which the state, $\vec{x} \in \mathbf{R}^n$, is a vector whose component x_i indicates the time at which a transition occurs on the i-th input wire. (The circuit is thought of as a closed system with the outputs fed back to the inputs. The dimension n is hence equal to the number of input wires.) A transition on a wire is a change in logic level from 1 to 0 or vice versa. The min-max function describes how the state changes; the parameters in the function being estimated from the internal gate delays in the circuit. If the circuit is in state $\vec{x}$ then $F(\vec{x})$ records the times at which the next transitions occur on the wires. It is easy to calculate the average time to the next transition over a sample of s state changes:

$$\frac{F^s(\vec{x}) - F^{s-1}(\vec{x}) + \cdots + F(\vec{x}) - \vec{x}}{s} = \frac{F^s(\vec{x}) - \vec{x}}{s}.$$

If we now let s tend to ∞ we see that the asymptotic average time to the next transition is given by the vector quantity

$$\lim_{s\to\infty} \frac{F^s(\vec{x})}{s}. \tag{3}$$

We refer to this asymptotic average, when the limit (3) exists, as the cycle time vector of the system. It appears to depend on $\vec{x}$ and hence to be a property not only of the system but also of where the system is started from. However, one of the first things that we shall learn about min-max functions is that (3), when it exists, is independent of the choice of $\vec{x}$ and is therefore characteristic of the system. This brings us to the main problem of this paper: when does the limit (3) exist and how can we calculate it when it does?

Before proceeding further we should clarify the implications for asynchronous circuits. Burns was the first to define a timing metric for asynchronous circuits, [2], and it can be shown that, under his assumptions, [2, §2.2.2], the cycle time vector as defined here has the simple form $(h, h, \cdots, h)$ where h is Burns's cycle period. The circuits considered by Burns are characterised by the fact that they only require maximum timing constraints. The corresponding min-max function F is max-only: $\wedge$ does not appear in any F_i. The theory of such functions is now well understood, as we shall discuss below. Burns was unable to give a rigorous definition of a timing metric for systems with both maximum and minimum constraints and he pointed out via an example, [2, §4.5], that his methods fail for such systems. Our work extends the theory to deal with this more general case.

We should also point out that min-max functions have been used to study timing problems in synchronous circuits as well, [16]. However, in this application area the critical problem is to calculate a fixed point of F: a vector $\vec{x}$ where $F(\vec{x}) = \vec{x}$. In related work we have given a necessary and sufficient condition for the existence of a fixed point for any min-max function, [6, Theorem 12]. In this paper we shall concentrate on the cycle time problem. A careful discussion of the applications to both asynchronous and synchronous timing problems is given in [11].

Having discussed how the application area naturally suggests the main problem of this paper, we shall step back slightly to explain how the theory of min-max functions fits into current theoretical research.

In concurrency theory, min-max functions can be thought of as describing systems which are timed analogues of untimed conflict free systems. To see why this is the case we note that a conflict free system may be identified with an {AND, OR} automaton, [9, §3]. In this event-based view, we are only concerned with whether or not an event occurs. When we add time to the picture, we need to know when the event will occur. It is clear that AND causality, in which an event has to wait for all the events in a set, translates naturally into a maximum timing constraint, while OR causality, in which an event has to wait only for the first event of a set, translates into a minimum timing constraint. A treatment along these lines appears in [10, §3].

The identification of conflict freedom with {AND, OR} causality is consistent with other well-known descriptions of conflict freedom such as persistence in Petri nets, [13], Keller's conditions for labelled transition systems, [12, §1.3], confluence in CCS, [14, Chapter 11], and conflict freedom in event structures, [18]. A precise statement of this appears in [8]. The mathematical theory of untimed conflict free systems is simple and has been known in essence for many years. There is only one theorem, which asserts that a certain poset is a semi-modular lattice, [8]. In contrast, the mathematical theory of timed conflict free systems, which is to say the theory of min-max functions, has already revealed some deep results and contains many open problems. We note in passing that it has been shown that timed marked graphs and, more generally, timed persistent Petri nets, have an interesting mathematical theory, [3], but the connection between this work and the theory of min-max functions has not yet been determined.

Of more immediate relevance to min-max functions is the work on max-plus algebra, [1, §3]. This is relatively unfamiliar in concurrency theory but deserves to be more widely known. It allows us to reduce the special case of a max-only function (or, dually, a min-only function) to a linear problem which can be dealt with by classical matrix methods. To see why this is so, note that, because $\vee$ is associative and commutative, any max-only function in the variables $x_1, \cdots, x_n$ can be reduced to the canonical form:

$$F_i(x_1, \cdots, x_n) = (a_{i1} + x_1 \vee \cdots \vee a_{in} + x_n), \tag{4}$$

where $a_{ij} \in \mathbf{R} \cup \{-\infty\}$. (The $-\infty$ merely serves as a zero for $\vee$; additive operations involving $-\infty$ work in the obvious way.) Now comes a very beautiful trick, which is due to Cuninghame-Green, [4]: redefine the operations on $\mathbf{R} \cup \{-\infty\}$ so that $+$ becomes $\vee$ and $\times$ becomes $+$. This new algebra is called max-plus algebra, [1, Definition 3.3]. The equation (4) can now be rewritten as

$$F(\vec{x}) = A\vec{x}$$

where $\vec{x}$ is considered as a column vector, $A = (a_{ij})$ is the matrix of parameter values and matrix multiplication is used on the right hand side.

The basic dynamical problem, of determining $F^s(\vec{x})$, now has an easy solution: it is simply the matrix product $A^s\vec{x}$ in max-plus algebra. More importantly, we can also show that the limit (3) exists and calculate it explicitly. We postpone further discussion of how to do this to §2. The theory of min-max functions can now be understood as a non-linear generalisation of max-plus algebra. The theory begins, in effect, where [1] ends.

We can now return to the cycle time problem which was enunciated above. What progress has been made towards solving it? In earlier work we identified and stated the Duality Conjecture, [6, §2], whose affirmative resolution would not only show that the limit (3) always exists but would also give a simple formula for calculating it. We also showed that if the Duality Conjecture were true then we could give a simple proof of Olsder's fixed point criterion for separated functions—the main result of [15]—as well as deduce several other powerful results. The Conjecture appears, therefore, to capture a fundamental property of min-max functions. The main result of the present paper, Theorem 3.1, is a proof of the Duality Conjecture for functions of dimension 2. This is the first rigorous evidence that the Conjecture is true in generality. Unfortunately, the proof does not extend to higher dimensions.

The rest of the paper falls into two parts. In the first part we state the elementary properties of min-max functions, recall the linear theory for max-only functions and state the Duality Conjecture. The second part gives the proof in dimension 2.

The author is very grateful to Professor Geert-Jan Olsder for his encouragement and hospitality during a visit to the University of Delft during which the details of the proof presented here were first worked out. The work described in this paper was undertaken as part of project STETSON, a joint project between Hewlett-Packard Laboratories and Stanford University on asynchronous hardware design.

2 Basic properties and the linear theory

In this section we study some of the elementary properties of min-max functions and show, in particular, why the limit (3) is independent of $\vec{x}$. We then discuss the special case of max-only functions using the linear methods of max-plus algebra. This leads us to the Duality Conjecture via the important concept of projections.

We shall frequently use numerical operations and relations and apply them to vectors. These should always be assumed to be applied to each component separately. Hence $\vec{u} \leq \vec{v}$ means $u_i \leq v_i$ for each i. Similarly, $(\bigwedge_l \vec{a}_l)_i = \bigwedge_l (\vec{a}_l)_i$. Let $\vec{c}(h) = (h, h, \cdots, h)$ denote the vector each of whose components has the same value h.

The first useful property of a min-max function of dimension n is that it is monotone:

$$\vec{u} \leq \vec{v} \implies F(\vec{u}) \leq F(\vec{v}). \qquad (5)$$

Second, F is homogeneous, in the sense that, for any $h \in \mathbf{R}$,

$$F(\vec{u} + \vec{c}(h)) = F(\vec{u}) + \vec{c}(h). \qquad (6)$$

This follows easily from (2), [7, Lemma 2.3]. The third and final elementary property is not quite so obvious. Let $|\vec{u}|$ denote the l^∞, or maximum, norm on vectors in $\mathbf{R}^n$: $|\vec{u}| = \bigvee_{1 \leq i \leq n} |u_i|$, where $|u_i|$ is the usual absolute value on real numbers. We require a simple preliminary observation.

Lemma 2.1 *For any real numbers $a_i, b_i \in \mathbf{R}$ with $1 \leq i \leq n$,*

$$|(\bigwedge_{1 \leq i \leq n} a_i) - (\bigwedge_{1 \leq i \leq n} b_i)| \;\leq\; \bigvee_{1 \leq i \leq n} |a_i - b_i|$$

and similarly with $\bigwedge$ replaced by $\bigvee$ on the left hand side.

Proof: Suppose that $a_k = \bigwedge_{1 \leq i \leq n} a_i$ and $b_j = \bigwedge_{1 \leq i \leq n} b_i$. If $a_k \leq b_j$ then $|a_k - b_j| \leq |a_k - b_k|$. If $b_j \leq a_k$ then $|a_k - b_j| \leq |a_j - b_j|$. Similarly with $\bigwedge$ replaced by $\bigvee$ on the left hand side.

QED

Lemma 2.2 *(Non-expansive property.) Let F be a min-max function of dimension n. If $\vec{u}, \vec{v} \in \mathbf{R}^n$ then $|F(\vec{u}) - F(\vec{v})| \leq |\vec{u} - \vec{v}|$.*

Proof: Assume that $\vec{u}$ and $\vec{v}$ are fixed and let $h = |\vec{u} - \vec{v}|$. It is sufficient to show that if f is any min-max expression of n variables, then $|f(\vec{u}) - f(\vec{v})| \leq h$. The proof of this is by induction on the structure of f. If $f \equiv x_k$, then $|f(\vec{u}) - f(\vec{v})| = |u_k - v_k| \leq h$ as required. Now suppose as an inductive hypothesis that the required result holds for f and g. It is obvious that it must then hold for $f + a$. By the preceding Lemma for $n = 2$ and the inductive hypothesis,

$$|(f \wedge g)(\vec{u}) - (f \wedge g)(\vec{v})| \leq |f(\vec{u}) - f(\vec{v})| \vee |g(\vec{u}) - g(\vec{v})| \leq h.$$

Similarly for $f \vee g$. The result follows by structural induction.

QED

It is important to note that F is not contractive. That is, there is no $0 < \lambda < 1$ such that $|F(\vec{u}) - F(\vec{v})| \leq \lambda|\vec{u} - \vec{v}|$. If there were, then the Contraction Mapping Theorem would imply that F had a unique fixed point and that $F^s(x)$ converges to it as $s \to \infty$. The dynamic behaviour of F is more complicated than that. However, suppose that the limit (3) exists at some point $\vec{x} \in \mathbf{R}^n$. Then it follows immediately from the non-expansive property that (3) exists everywhere in $\mathbf{R}^n$ and has the same value.

Definition 2.1 *Let F be a min-max function. If the limit (3) exists somewhere, it is called the cycle time vector of F and denoted by $\vec{\chi}(F) \in \mathbf{R}^n$.*

When does the cycle time vector exist? Let us begin to answer this by considering the special case of a max-only function. Suppose that F is a max-only function of dimension n and that A is the associated $n \times n$ matrix in max-plus algebra. For example, the following max-only function of dimension 3

$$\begin{aligned} F_1(x_1, x_2, x_3) &= x_2 + 2 \vee x_3 + 5 \\ F_2(x_1, x_2, x_3) &= x_2 + 1 \\ F_3(x_1, x_2, x_3) &= x_1 - 1 \vee x_2 + 3 \end{aligned} \tag{7}$$

has the associated max-plus matrix

$$\begin{pmatrix} -\infty & 2 & 5 \\ -\infty & 1 & -\infty \\ -1 & 3 & -\infty \end{pmatrix}. \tag{8}$$

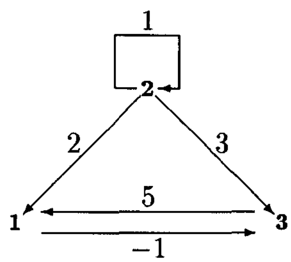
We recall that the precedence graph of A, [1, Definition 2.8], denoted $\mathcal{G}(A)$, is the directed graph with annotated edges which has nodes $\{1, 2, \cdots, n\}$ and an edge from j to i if, and only if, $A_{ij} \neq -\infty$. The annotation on this edge is then the real number A_{ij}. We shall denote an edge from j to i by $i \leftarrow j$. A path in this graph has the usual meaning of a chain of directed edges: a path from i_m to i_1 is a sequence of nodes $i_1, \cdots, i_m$ such that $i_j \leftarrow i_{j+1}$ for $1 \leq j < m$. A circuit is a path which starts and ends at the same node: $i_1 = i_m$. This includes the possibility that $m = 1$. A circuit is elementary if the nodes $i_1, \cdots, i_{m-1}$ are all distinct. A path or circuit is upstream from node i if there is a path in $\mathcal{G}(A)$ from some node on the path or circuit to node i. The weight of a path p, $|p|_{\mathsf{w}}$, is the sum of the annotations on the edges in the path:

$$|p|_{\mathsf{w}} = \sum_{j=1}^{m-1} A_{i_j i_{j+1}}.$$

The length of a path, $|p|_\ell$, is the number of edges in the path: $|p|_\ell = m-1$. If g is a circuit, the ratio $|g|_{\mathsf{w}}/|g|_\ell$ is the cycle mean of the circuit, [1, Definition 2.18].

Definition 2.2 *If A is an $n \times n$ matrix in max-plus algebra, let $\vec{\mu}(A) \in (\mathbf{R} \cup \{-\infty\})^n$ be the vector such that $\mu_i(A) = \bigvee\{\ |g|_{\mathsf{w}}/|g|_\ell \mid g$ a circuit in $\mathcal{G}(A)$ upstream from node $i\}$.*

It is not difficult to check that in calculating $\mu_i(A)$ it is only necessary to consider elementary circuits, of which there are only finitely many. By convention, $\bigvee \emptyset = -\infty$. Because each component expression of a max-only function must be non-empty, it follows that each row of the associated max-plus matrix must have an entry not equal to $-\infty$. There is, therefore, always some circuit upstream from any node. Hence, for any matrix associated to a max-only function, $\vec{\mu}(A) \in \mathbf{R}^n$. The precedence graph of example (8) is shown below



and the reader should have no difficulty in showing that $\vec{\mu}(A) = (2,1,2)$. The significance of $\vec{\mu}(A)$ is revealed by the following result whose proof may be found in [6].

Proposition 2.1 *([6, Proposition 2.1]) Let F be a max-only function and A the associated matrix in max-plus algebra. The limit (3) exists and $\vec{\chi}(F) = \vec{\mu}(A)$.*

We now want to consider a general min-max function. Before doing so we need to make some remarks about duality. Let f be a min-max expression. By similar arguments to those used to find the canonical form (4) it is easy to see that f can be placed in conjunctive form:

$$f = f_1 \wedge f_2 \wedge \cdots \wedge f_k,$$

where the f_i are distinct max-only expressions in canonical form. (This expression can be made unique, up to permutation of the f_i, by throwing out redundant terms, [7, Theorem 2.1], although one has to be careful to define "redundant" correctly. The resulting expression is called conjunctive normal form. Although normal forms were used to state the Duality Conjecture in [6] it is not essential to use them and we shall not do so in this paper.) By a dual argument, we can also put f into disjunctive form:

$$f = g_1 \vee g_2 \vee \cdots \vee g_l,$$

where the g_i are distinct min-only expressions in canonical form. Note that expressions in conjunctive form have parameters in $\mathbf{R} \cup \{-\infty\}$ while those in disjunctive form have parameters in $\mathbf{R} \cup \{+\infty\}$.

There is a simple algorithm for moving back and forth between conjunctive and disjunctive form which we shall need to use in §3. We explain it here by working through an example. Consider the min-max expression of 2 variables,

$$f = (a + x_1 \vee b + x_2) \wedge c + x_1, \tag{9}$$

where $a, b, c \in \mathbf{R}$. This is effectively in conjunctive form but to be more precise we should write f as

$$(a + x_1 \vee b + x_2) \wedge (c + x_1 \vee -\infty + x_2).$$

To express f in disjunctive form we go back to the initial min-max expression (9) and rewrite each individual term $a_i + x_i$ (where $a_i \neq -\infty$) in disjunctive form. This gives

$$((a + x_1 \wedge +\infty + x_2) \vee (+\infty + x_1 \wedge b + x_2)) \wedge (c + x_1 \wedge +\infty + x_2).$$

We now use the distributivity of $\wedge$ over $\vee$ to interchange the order of the two operations and get

$$((a \wedge c) + x_1 \wedge +\infty + x_2) \vee (c + x_1 \wedge b + x_2), \tag{10}$$

which is in disjunctive form.

Now suppose that F is an arbitrary min-max function of dimension n. Each component of F can be placed in conjunctive form as above:

$$F_k(\vec{x}) = (A^k_{11} + x_1 \vee \cdots \vee A^k_{1n} + x_n) \wedge \cdots \wedge (A^k_{\ell(k)1} + x_1 \vee \cdots \vee A^k_{\ell(k)n} + x_n), \tag{11}$$

where $A^k_{ij} \in \mathbf{R} \cup \{-\infty\}$. Here $\ell(k)$ is the number of conjunctions in the component F_k. We can now associate a max-plus matrix A to F by choosing, for the k-th row of the matrix, one of the $\ell(k)$ conjunctions in (11): $A_{kj} = A^k_{i_k j}$ where $1 \leq i_k \leq \ell(k)$ specifies which conjunction is chosen in row k.

Definition 2.3 *The matrix A constructed in this way is called a max-only projection of F. A set of max-only projections is the collection of all such matrices from a single conjunctive form such as (11). Dually, a set of min-only projections is constructed from a disjunctive form.*

Sets of max-only projections are not unique. However, if we use conjunctive normal form then it follows from [7, Theorem 2.1] that the corresponding set of normal max-only projections is uniquely defined for any function F. Sets of projections can be quite large: the function (11) has $\prod_{1 \leq i \leq n} \ell(i)$ distinct max-only projections.

At this point an example may be helpful. Consider the min-max function of dimension 2:

$$\begin{array}{rcl} F_1(x_1, x_2) & = & (a + x_1 \vee b + x_2) \wedge c + x_1 \\ F_2(x_1, x_2) & = & (t + x_1 \wedge u + x_2) \end{array} \tag{12}$$

where $a, b, c, t, u \in \mathbf{R}$. F_1 is already in conjunctive form while F_2 is in disjunctive form. We first put F_2 into conjunctive form using the algorithm discussed above:

$$F_2(x_1, x_2) = (t + x_1 \vee -\infty + x_2) \wedge (-\infty + x_1 \vee u + x_2)$$

and then read off a set of max-only projections:

$$\left\{ \begin{pmatrix} a & b \\ t & -\infty \end{pmatrix}, \begin{pmatrix} a & b \\ -\infty & u \end{pmatrix}, \begin{pmatrix} c & -\infty \\ t & -\infty \end{pmatrix}, \begin{pmatrix} c & -\infty \\ -\infty & u \end{pmatrix} \right\}. \tag{13}$$

Dually, we can put F_1 into disjunctive form—an exercise already performed in (10):

$$F_1(x_1, x_2) = ((a \wedge c) + x_1 \wedge +\infty + x_2) \vee (c + x_1 \wedge b + x_2)$$

and read off a set of min-only projections:

$$\left\{ \begin{pmatrix} a \wedge c & +\infty \\ t & u \end{pmatrix}, \begin{pmatrix} c & b \\ t & u \end{pmatrix} \right\}. \tag{14}$$

We hope this has clarified these important constructs. For the remainder of the paper we shall use the letter A for max-plus matrices and the letter B for min-plus matrices. It will also be convenient to introduce the notation $\vec{\eta}(B)$ to indicate the vector which is "dual" to $\vec{\mu}(A)$. In other words, $\vec{\eta}(B)$ is the vector of minimum upstream cycle means in $\mathcal{G}(B)$.

Let F be any min-max function of dimension n and let P and Q be sets of max-only and min-only projections, respectively, of F. It is clear from the construction above that, for any $A \in P$, $F(\vec{x}) \leq A\vec{x}$ for all $\vec{x} \in \mathbf{R}^n$. It follows from (5) that $F^s(\vec{x}) \leq A^s\vec{x}$ for all $s \geq 0$. Now choose $\epsilon > 0$. It then follows from Proposition 2.1 that, for all sufficiently large s, $F^s(\vec{x})/s \leq \vec{\mu}(A) + \vec{c}(\epsilon)$. Since this holds for any max-only projection in P, and there are only finitely many such, we see that $F^s(\vec{x})/s \leq (\bigwedge_{A \in P} \vec{\mu}(A)) + \vec{c}(\epsilon)$ for all sufficiently large s. By a dual argument applied to the min-only projections of F, we can conclude that

$$(\bigvee_{B \in Q} \vec{\eta}(B)) - \vec{c}(\epsilon) \leq \frac{F^s(\vec{x})}{s} \leq (\bigwedge_{A \in P} \vec{\mu}(A)) + \vec{c}(\epsilon). \tag{15}$$

for all sufficiently large s.

Conjecture 2.1 *(The Duality Conjecture.) Let F be any min-max function and let P and Q be any sets of max-only and min-only projections, respectively, of F. Then,*

$$\bigvee_{B \in Q} \vec{\eta}(B) = \bigwedge_{A \in P} \vec{\mu}(A). \tag{16}$$

The significance of this should be clear. It implies that any min-max function has a cycle time and gives us a formula for computing it. It is worth working through an example to see how the numbers come out. We reproduce below the details for example (12) whose projections were worked out above. The $\vec{\mu}(A)$ of the max-only projections of (12), in the order in which they are listed in (13), are shown below as column vectors:

$$\begin{pmatrix} a \vee (b+t)/2 \\ a \vee (b+t)/2 \end{pmatrix}, \begin{pmatrix} a \vee u \\ u \end{pmatrix}, \begin{pmatrix} c \\ c \end{pmatrix}, \begin{pmatrix} c \\ u \end{pmatrix}.$$

So the right hand side of (16) is

$$\begin{pmatrix} (a \vee (b+t)/2) \wedge (a \vee u) \wedge c \\ (a \vee (b+t)/2) \wedge u \wedge c \end{pmatrix}. \tag{17}$$

The $\vec{\eta}(B)$ of the min-only projections, in the same order as they appear in (14), are:

$$\begin{pmatrix} a \wedge c \\ u \wedge a \wedge c \end{pmatrix}, \begin{pmatrix} c \wedge (b+t)/2 \wedge u \\ c \wedge (b+t)/2 \wedge u \end{pmatrix},$$

and so the left hand side of (16) is

$$\begin{pmatrix} (a \wedge c) \vee (c \wedge (b+t)/2 \wedge u) \\ (u \wedge a \wedge c) \vee (c \wedge (b+t)/2 \wedge u) \end{pmatrix}. \tag{18}$$

The reader will have no trouble confirming that (17) and (18) are identical. The calculation gives little hint as to why the numbers come out to be the same. We shall try and understand this in the next section. Before moving on, we note that since ϵ was arbitrary, (15) already tells us that

$$\bigvee_{B\in Q} \vec{\eta}(B) \leq \bigwedge_{A\in P} \vec{\mu}(A). \tag{19}$$

The next section is devoted to showing that the reverse inequality also holds, at least when F has dimension 2.

3 The Duality Conjecture in dimension 2

Theorem 3.1 *If F is any min-max function of dimension 2 then $\bigvee_{B\in Q} \vec{\eta}(B) = \bigwedge_{A\in P} \vec{\mu}(A)$.*

The proof of this occupies the entire section. We begin with some preparatory remarks and notation, which are tailored to dimension 2. If A is a 2×2 matrix in max-plus algebra,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with $a, b, c, d \in \mathbf{R} \cup \{-\infty\}$, then the maximum upstream cycle mean of component 1 is

$$\mu_1(A) = \begin{cases} a \vee (b+c)/2 \vee d & \text{if } b \neq -\infty \\ a & \text{otherwise} \end{cases}. \tag{20}$$

This equation is valid even when a, b, c, d take on the value $-\infty$; the point of the second formula being that, if $b = -\infty$, then node 2 is no longer upstream of node 1. In all other cases, the first formula gives the correct answer. A similar dual equation holds for $\eta_1(B)$.

Now suppose that we have some expression of the form $e = \bigwedge_{i\in\mathcal{L}}(a_i \vee b_i \vee c_i)$ where $\mathcal{L}$ is some finite index set and $a_i, b_i, c_i \in \mathbf{R} \cup \{-\infty\}$. In what follows we shall want to rewrite expressions like this "the other way round", as a maxima of minima. Let $\mathsf{P}^n(\mathcal{L})$ denote the set of partitions of $\mathcal{L}$ into n disjoint pieces:

$$\mathsf{P}^n(\mathcal{L}) = \{\{U_1, \cdots, U_n\} \mid U_i \subseteq \mathcal{L},\ \mathcal{L} = U_1 \cup \cdots \cup U_n \text{ and } U_i \cap U_j = \emptyset \text{ for } i \neq j\}.$$

By distributivity, it follows easily that

$$e = \bigvee_{\{S,T,U\}\in\mathsf{P}^3(\mathcal{L})} \bigwedge_{i\in S} a_i \wedge \bigwedge_{i\in T} b_i \wedge \bigwedge_{i\in U} c_i.$$

In this equation we rely on the convention that $\bigwedge \emptyset = +\infty$. If both $-\infty$ and $+\infty$ appear in one of the partition terms (for instance, if S contains an index i with $a_i = -\infty$ and $U = \emptyset$) then the $-\infty$ will "win" and the corresponding partition will not contribute to the final maximum.

Let F be a min-max function of dimension 2 and let us use x and y in place of the variables x_1 and x_2. We may write F in conjunctive form in the following way:

$$\begin{array}{rcl} F_1(x,y) & = & \bigwedge_{1\leq i\leq n_1}(a_i + x \vee b_i + y) \\ F_2(x,y) & = & \bigwedge_{1\leq j\leq n_2}(c_j + x \vee d_j + y), \end{array} \tag{21}$$

where $a_i, b_i, c_j, d_j \in \mathbf{R} \cup \{-\infty\}$. It will be convenient to identify the sets $I_1, I_2 \subseteq \{1, \cdots, n_1\}$ and $J_1, J_2 \subseteq \{1, \cdots, n_2\}$ where

$$\begin{array}{rclrcl} I_1 & = & \{i \mid b_i = -\infty\} & I_2 & = & \{i \mid a_i = -\infty\} \\ J_1 & = & \{j \mid d_j = -\infty\} & J_2 & = & \{j \mid c_j = -\infty\}. \end{array}$$

Because not both a_i and b_i can be $-\infty$, and similarly for c_j and d_j, these sets satisfy the restrictions

$$I_1 \cap I_2 = \emptyset \text{ and } J_1 \cap J_2 = \emptyset. \tag{22}$$

This method of indexing F is convenient because it allows us to write down the max-only projections quite simply as:

$$A(i,j) = \left(\begin{array}{cc} a_i & b_i \\ c_j & d_j \end{array} \right),$$

where (i,j) runs over the index set $\{1,\cdots,n_1\} \times \{1,\cdots,n_2\}$. (It is useful to picture this index set as a rectangular region of lattice points in the plane. Most of the assertions that we shall make regarding subsets of it can be easily visualised in this way.) Let P denote the corresponding set of max-only projections. If we apply the dualising algorithm used on example (9) to construct a disjunctive form for F then we can read off a set of min-only projections. Let Q denote this set. Let R be the first component of the right hand side of (16): $R = \bigwedge_{A\in P} \mu_1(A)$. Similarly, let L be the first component of the left hand side: $L = \bigvee_{B\in Q} \eta_1(B)$. We established in (19) that $L \leq R$. We shall now show that $R \leq L$. It is clear that if we can do this for the first component of any min-max function of dimension 2 then we will have proved Theorem 3.1. A few final pieces of notation are required: if $X \subseteq A \times B$, then $\pi_1 X$ and $\pi_2 X$ will denote the projections of X onto A and B, respectively. If X, A are sets then $A \backslash X = \{a \in A \mid a \notin X\}$. If $X \subseteq A$ then $\overline{X} = X \backslash A$ when A is clear from the context.

We can now begin the proof in earnest. Using (20) we can rewrite R "the other way round" by the method discussed above. If $\{S,T,U\} \in \mathsf{P}^3(\{1,\cdots,n_1\} \times \{1,\cdots,n_2\})$ it will be convenient to introduce the auxiliary function

$$\rho(S,T,U) = \bigwedge_{p\in\pi_1 S} a_p \wedge \bigwedge_{(q,r)\in T} \frac{b_q + c_r}{2} \wedge \bigwedge_{u\in\pi_2 U} d_u. \tag{23}$$

We may then write

$$R = \bigvee_{\{S,T,U\}\in\mathsf{P}^3(\{1,\cdots,n_1\} \times \{1,\cdots,n_2\})} \rho(S,T,U) \tag{24}$$

where we must assume that $\pi_1 T \cap I_1 = \pi_1 U \cap I_1 = \emptyset$ because of (20). We may further disregard those partitions which make a contribution of $-\infty$ to the maximum. So we may finally assume that the partitions in (24) satisfy the restrictions

$$\begin{array}{rcccl} \pi_1 S \cap I_2 & = & \emptyset & & \\ \pi_1 T \cap I_1 & = & \pi_2 T \cap J_2 & = & \emptyset \\ \pi_2 U \cap J_1 & = & \pi_1 U \cap I_1 & = & \emptyset. \end{array} \tag{25}$$

It is easy to deduce from these equations and the fact that $\{S,T,U\}$ is a partition, that $I_1 \times \{1,\cdots,n_1\} \subseteq S$. In particular,

$$I_1 \subseteq \pi_1 S. \tag{26}$$

We want to compare (24) with L. To compute L we need the set Q of min-only projections which are obtained by dualising (21). The dualisation amounts to writing the component expressions of F "the other way round". We leave it to the reader to check that the resulting min-only projections can be indexed as

$$B(X,Y) = \left(\begin{array}{cc} \bigwedge_{p\in X} a_p & \bigwedge_{q\in \overline{X}} b_q \\ \bigwedge_{r\in Y} c_r & \bigwedge_{u\in \overline{Y}} d_u \end{array} \right), \tag{27}$$

where $X \subseteq \{1, \cdots, n_1\}$ and $Y \subseteq \{1, \cdots, n_2\}$ are any subsets satisfying the restrictions:

$$\begin{array}{lllll} I_1 & \subseteq & X & \subseteq & \{1, \cdots, n_1\} \backslash I_2 \\ J_1 & \subseteq & Y & \subseteq & \{1, \cdots, n_2\} \backslash J_2. \end{array} \tag{28}$$

Min-only matrices must have entries in $\mathbf{R} \cup \{+\infty\}$ and it is clear that the restrictions in (28) will guarantee this. It is largely for this purpose that the subsets I_1, I_2 and J_1, J_2 were introduced.

Lemma 3.1 *With the details above, suppose that for each partition $\{S, T, U\}$ satisfying the restrictions in (25), it is possible to find X, Y satisfying the restrictions in (28) such that*

$$\rho(S, T, U) \leq \eta_1 B(X, Y).$$

Then $R \leq L$.

Proof: With the restrictions in (25) and (28), we have

$$R = \bigvee_{S,T,U} \rho(S, T, U) \leq \bigvee_{X,Y} \eta_1 B(X, Y) = L.$$

QED

All we have done so far is book-keeping. The main part of the argument is still to come. The crux of the proof hinges on the nature of the set T. If we think about the form of $\eta_1 B(X, Y)$, as calculated from (the dual version of) (20), then the part played by T is the set $\overline{X} \times Y$. This differs from T in being a product subset, or rectangle, in $\{1, \cdots, n_1\} \times \{1, \cdots, n_2\}$. It is this clue which gives rise to the argument which follows. The idea is to replace the partition (S, T, U) by a new partition (S', T', U') which still satisfies (25) but for which $\rho(S, T, U) \leq \rho(S', T', U')$. The new partition will have T' rectangular. Such partitions can be dealt with relatively simply. Rectangularisation is thus the crucial step.

Choose some partition $\{S, T, U\}$ satisfying (25). We first need to deal with the possibility that $T = \emptyset$. Assume that this is so. It follows from (23) that

$$\rho(S, \emptyset, U) = \bigwedge_{p \in \pi_1 S} a_p \wedge \bigwedge_{u \in \pi_2 U} d_u.$$

Now suppose further that $\pi_2 U = \{1, \cdots, n_2\}$. It follows from (25) that $J_1 = \emptyset$. Let $X = \pi_1(S)$ and $Y = \emptyset$. It follows from (25) and (26) that X, Y satisfy (28). Since $Y = \emptyset$, the corresponding matrix $B(X, Y)$, shown in (27), has $+\infty$ in the bottom left corner. If $\overline{X} \neq \emptyset$, it follows from (20) that,

$$\eta_1 B(X, Y) = \bigwedge_{p \in X} a_p \wedge \bigwedge_{u \in \overline{Y}} d_u$$

and it is clear that $\rho(S, \emptyset, U) = \eta_1 B(X, Y)$. If $\overline{X} = \emptyset$ then certainly $\rho(S, \emptyset, U) \leq \eta_1 B(X, Y)$ since the latter omits the contribution from U. In either case we are done. Now suppose that $\pi_1 S \neq \{1, \cdots, n_1\}$. It then follows from (25) that $\pi_2 U = \{1, \cdots, n_2\}$ (this is where a picture comes in handy) and we have already done this case. So we may assume that $\pi_1 S = \{1, \cdots, n_1\}$ and hence that $I_2 = \emptyset$. Let $X = \{1, \cdots, n_1\}$, which certainly satisfies (28), and let Y be any subset of $\{1, \cdots, n_2\}$ which satisfies (28). We can always choose such a subset in view of (22). Because $\overline{X} = \emptyset$, it follows from (20) that $\eta_1 B(X, Y) = \bigwedge_{p \in X} a_p$. Hence $\rho(S, \emptyset, U) \leq \eta_1 B(X, Y)$ and once again we are done. This deals with all the possibilities when $T = \emptyset$.

Now assume that $T \neq \emptyset$. Suppose that we can find $u \in \pi_1 T$ and $v \in \pi_2 T$ such that $(u, v) \notin T$. Then either $(u, v) \in S$ or $(u, v) \in U$. If $(u, v) \in S$ then let $D \subseteq \{1, \cdots, n_1\} \times \{1, \cdots, n_2\}$ be the set $D = \{x \in T \mid \pi_1 x = u\}$. Evidently, $D \neq \emptyset$. Construct a new partition $\{S', T', U'\}$ such that $S' = S \cup D$, $T' = T \setminus D$ and $U' = U$. It is clear that this is still a partition of $\{1, \cdots, n_1\} \times \{1, \cdots, n_2\}$. We need to check that it satisfies (25). Since T has got smaller, it follows that T' cannot violate (25) and, of course, U has not changed. As for S, it is easy to see that $\pi_1 S' = \pi_1 S$, so that S also satisfies (25). We thus have a good partition. Furthermore, since T' has got smaller while $\pi_1 S' = \pi_1 S$ and $\pi_2 U' = \pi_2 U$, it is easy to see that $\rho(S, T, U) \leq \rho(S', T', U')$. If $(u, v) \in U$ then we move elements from T to U and a similar argument works. We can now carry on constructing new partitions in this way. Since T is finite and strictly decreases each time, the process can only stop in two ways. Either we end up with $T = \emptyset$, which we have already dealt with, or we find that we can no longer choose (u, v) satisfying the requirements above. But it must then be the case that $T = \pi_1 T \times \pi_2 T$. Hence we may assume that T is non-empty and rectangular. The importance of this stems from the following elementary fact.

Lemma 3.2 *With the above details, if T is rectangular, then*

$$\bigwedge_{(q,r) \in T} (b_q + c_r)/2 = ((\bigwedge_{q \in \pi_1 T} b_q) + (\bigwedge_{r \in \pi_2 T} c_r))/2.$$

Proof: The rectangularity of T implies that

$$\bigwedge_{(q,r) \in T} (b_q + c_r)/2 \quad = \quad \bigwedge_{q \in \pi_1 T} \bigwedge_{r \in \pi_2 T} (b_q + c_r)/2.$$

We can now use (2) twice to rewrite this as follows:

$$\begin{aligned} &= \bigwedge_{q \in \pi_1 T} (b_q/2 + (\bigwedge_{r \in \pi_1 T} c_r/2)) \\ &= ((\bigwedge_{q \in \pi_1 T} b_q) + (\bigwedge_{r \in \pi_2 T} c_r))/2. \end{aligned}$$

QED

The remainder of the argument resembles the case when $T = \emptyset$. Suppose first that $\pi_2 U = \{1, \cdots, n_2\}$ so that $J_1 = \emptyset$. Let $X = \pi_1 S$ and $Y = \emptyset$. As before, these satisfy (28). The corresponding $B(X, Y)$ has $+\infty$ in the bottom left corner. It follows from (20) that $\rho(S, T, U) \leq \eta_1 B(X, Y)$ since the latter simply omits the contribution coming from T, if $\overline{X} \neq \emptyset$, and from both T and U, if $\overline{X} = \emptyset$. Now suppose that $\pi_1 S = \{1, \cdots, n_1\}$ so that $I_2 = \emptyset$. Let $X = \{1, \cdots, n_1\}$, which certainly satisfies (28), and choose any Y which also satisfies (28), which we may always do by (22). The corresponding $B(X, Y)$ has $+\infty$ in the top right corner. It follows from (20) that $\rho(S, T, U) \leq \eta_1 B(X, Y)$ since the latter omits the contributions from both T and U. Now let $\overline{X} = \pi_1 T$. If $X \not\subseteq \pi_1 S$ then it follows from (25) that $\pi_2 U = \{1, \cdots, n_2\}$, which we have already considered. So we may assume that $X \subseteq \pi_1 S$ and so $\bigwedge_{p \in \pi_1 S} a_p \leq \bigwedge_{p \in X} a_p$. Furthermore, it is easy to see that X satisfies (28). Let $Y = \pi_2 T$ and suppose that $\overline{Y} \not\subseteq \pi_2 U$. Then, in a similar way, it must be the case that $\pi_1 S = \{1, \cdots, n_1\}$, which we have also considered. Hence, we may also assume that $\overline{Y} \subseteq \pi_2 U$ and so $\bigwedge_{u \in \pi_2 U} d_u \leq \bigwedge_{u \in \overline{Y}} d_u$. Furthermore, Y also satisfies (28). But now,

$\rho(S,T,U) \leq \eta_1 B(X,Y)$ because in the latter the contribution from $\overline{X} \times Y$ is equal to that from T by Lemma 3.2, while the other contributions have got larger. This completes the proof of Theorem 3.1.

The argument we have presented is straightforward once the details of the book-keeping have been mastered. A similar approach can be attempted in higher dimensions, albeit at the cost of vastly increased book-keeping. It is not the book-keeping that defeats this, however. It turns out that Lemma 3.1 is no longer of any use. There is an example in dimension 3 such that, for a given partition of the form $\{S,T,U\}$ (but now requiring 8 entries), there is no single min-plus matrix satisfying the conditions of Lemma 3.1. Different min-plus matrices are required for different values of the parameters in F. This does not happen in dimension 2 as we have just seen. Attempting to force through a proof along these lines runs into a wall of technical difficulties in higher dimensions. It seems clear that some new ideas are required to make further progress on the Duality Conjecture.

4 Conclusion

The dynamic approach that we have sketched here has already succeeded in calculating numerical quantities of importance in engineering applications. It also presents an attractive collection of mathematical problems which are easy to formulate but difficult to prove. We have studied one such problem, perhaps the most fundamental one, in this paper. Other open problems are discussed in [6, 7].

It is interesting to speculate on how the theory could be extended to deal with a broader class of reactive systems. It is well known that digital circuits such as latches exhibit a variety of oscillatory behaviour (for example, metastability). These are outside the repertoire of min-max functions. It is also well known that the next stage of behavioural complexity in dynamical systems is non-periodic behaviour such as chaos, [5]. Is it possible to model metastability by chaos in a suitable dynamical system? Because of the great progress that has been made towards developing numerical measures of chaos (Liapunov exponents, entropy, etc), this could open up a new chapter in the study of metastability. In order to explore some of these possibilities we have tried to exploit the analogy between untimed systems with $\{\mathsf{AND}, \mathsf{OR}\}$ causality and timed systems described by min-max functions. By extending the untimed models to include NOT, or negation, [8, 9], we are hoping to understand how the theory of min-max functions may be extended to deal with timed conflict. There are many difficulties with developing a consistent theory along these lines, [8], and it still remains to be seen whether this will provide a foundation for a dynamic theory in the presence of conflict.

References

[1] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronization and Linearity.* Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, 1992.

[2] S. M. Burns. *Performance Analysis and Optimization of Asynchronous Circuits.* PhD thesis, California Institute of Technology, 1990.

[3] J. Campos, G. Chiola, and M. Silva. Ergodicity and throughput bounds of Petri nets with unique consistent firing count vector. *IEEE Transactions on Software Engineering*, 17(2):117–125, 1991.

[4] R. A. Cuninghame-Green. Describing industrial processes with interference and approximating their steady-state behaviour. *Operational Research Quarterly*, 13(1):95–100, 1962.

[5] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems.* Addison-Wesley Publishing Company, Inc., 1989.

[6] J. Gunawardena. Cycle times and fixed points of min-max functions. To appear in the Proceedings of the *11th International Conference on Analysis and Optimization of Systems*, Sophia-Antipolis, France, June 1994.

[7] J. Gunawardena. Min-max functions. Submitted to *Discrete Event Dynamic Systems.*

[8] J. Gunawardena. The Muller unfolding. In preparation.

[9] J. Gunawardena. A generalized event structure for the Muller unfolding of a safe net. In E. Best, editor, *CONCUR'93 - 4th International Conference on Concurrency Theory*, pages 278–292. Springer LNCS 715, 1993.

[10] J. Gunawardena. Periodic behaviour in timed systems with {AND, OR} causality. Part I: systems of dimension 1 and 2. Technical Report STAN-CS-93-1462, Department of Computer Science, Stanford University, February 1993.

[11] J. Gunawardena. Timing analysis of digital circuits and the theory of min-max functions. In *TAU'93, ACM International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, September 1993.

[12] R. M. Keller. A fundamental theorem of asynchronous parallel computation. In T. Y. Feng, editor, *Parallel Processing*, pages 102–112. Springer LNCS 24, 1975.

[13] L. H. Landweber and E. L. Robertson. Properties of conflict-free and persistent Petri nets. *Journal ACM*, 25(3):352–364, 1978.

[14] R. Milner. *Communication and Concurrency.* International Series in Computer Science. Prentice-Hall, 1989.

[15] G. J. Olsder. Eigenvalues of dynamic max-min systems. *Discrete Event Dynamic Systems*, 1:177–207, 1991.

[16] T. Szymanski and N. Shenoy. Verifying clock schedules. In *Digest of Technical Papers of the IEEE International Conference on Computer-Aided Design of Integrated Circuits*, pages 124–131. IEEE Computer Society, 1992.

[17] J. Vytopil, editor. *Formal Techniques in Real-Time and Fault-Tolerant Systems.* Springer LNCS 571, 1991.

[18] G. Winskel. Event structures. In W. Brauer, W. Reisig, and G. Rozenberg, editors, *Advances in Petri Nets.* Springer LNCS 255, 1987.

Note to the reader

Copies of [6]-[11] are available by ftp from **hplose.hpl.hp.com**, IP address 15.254.100.100. Use "anonymous" as user name and refer to "/pub/jhcg/README" for more information.