Chapter 1

Idempotent Analysis

1.0. Introduction

The first chapter deals with idempotent analysis per se. To make the presentation self-contained, in the first two sections we define idempotent semirings, give a concise exposition of idempotent linear algebra, and survey some of its applications. Idempotent linear algebra studies the properties of the semimodules A^n, $n \in \mathbb{N}$, over a semiring A with idempotent addition; in other words, it studies systems of equations that are linear in an idempotent semiring. Probably the first interesting and nontrivial idempotent semiring, namely, that of all languages over a finite alphabet, as well as linear equations in this semiring, was examined by S. Kleene [**?**] in 1956. This noncommutative semiring was used in applications to compiling and parsing (see also [**?**]). Presently, the literature on idempotent algebra and its applications to theoretical computer science (linguistic problems, finite automata, discrete event systems, and Petri nets), biomathematics, logic, mathematical physics, mathematical economics, and optimization, is immense; e.g., see [**?**, **?**]. In §1.2 we present the most important facts of the idempotent algebra formalism.

The semimodules A^n are idempotent analogs of the finite-dimensional vector spaces $\mathbb{R}^n$, and hence endomorphisms of these semimodules can naturally be called (idempotent) linear operators on A^n. These operators themselves obviously form a semimodule, which can be identified with the semimodule of A-valued $n \times n$ matrices. This module is a semialgebra with respect to the composition of mappings and will be denoted by $\mathrm{End}(A^n)$. Idempotent linear algebra underlies all computational schemes for discrete optimization on graphs, since the general discrete Bellman equation for the optimal value S of the performance functional (the length of the shortest path in a graph, the cardinality of the maximal matching, etc.) is just a linear equation in A^n for some semiring A with idempotent addition, that is, has the form

$$S = HS \oplus F,$$

where $F \in A^n$ and $H \in \mathrm{End}\, A^n$ are given and $S \in A^n$ is to be found. Thus, algorithms of discrete mathematics are methods for solving linear equations in A^n. Furthermore, it turns out that to each such equation there corresponds an object known as a *discrete computational medium* (this notion generalizes that of directed graph with given arc weights). Consequently, each discrete optimization problem can be represented by either a linear equation in A^n or a discrete computational medium. These two universal representations give rise to two general approaches to constructing universal convergence criteria for discrete algorithms. The first approach is algebraic; that is, the criteria are stated in terms of the matrix $H \in \mathrm{End}\, A^n$ that occurs in the equation (for example, in terms of the main invariants of H, such as the idempotent analogs of trace and spectrum). The second approach is geometric; that is, the criteria are stated in terms of the geometry of the corresponding computational medium. The proof of these criteria is based on general formulas for solutions of linear equations in a semiring. These formulas are idempotent analogs of the Duhamel and the Green formulas and of the path integral representation of solutions of usual linear equations.

In the standard arithmetic, the multiplication of a number a by a positive integer $m \in \mathbb{N}$ is equivalent to adding a to itself $m-1$ times. This obvious consideration implies that continuous additive functions on $\mathbb{R}^n$ are linear. The argument fails for idempotent semimodules A^n, and an additive mapping $A^n \to A$ need not be linear. However, some important discrete optimization problems, for example, nonclassical trajectory-type problems (with variable traversing time) are described by equations that are not linear but only additive (with respect to the idempotent operation $\oplus$). A trick proposed in [**?**] permits one to reduce additive problems to linear ones (but in the more complicated semiring of endomorphisms of an idempotent semigroup), which are easier to analyze. Algorithms more efficient than those known previously were developed on the basis of this reduction for trajectory-type problems and for related operation problems for flexible computer-aided manufacturing systems. The simple mathematical fact underlying this trick is that the semiring $\operatorname{End} M^n$ of endomorphisms of the semigroup M^n, where M is an idempotent semigroup (a set equipped with an idempotent semigroup operation), is isomorphic to the semiring of endomorphisms of the semimodule $(\operatorname{End} M)^n$, that is, linear operators on this semimodule.

After discussing idempotent algebra, we give a brief survey of typical problems and present the scheme of a general algorithm for solving linear problems in semirings, namely, the generalized wave (or Ford–Lee) algorithm. As we have already mentioned, the material of §1.2 is rather standard. We finish our exposition of idempotent linear algebra with a very brief review of recent developments.

In §§1.3–1.4 we start dealing with idempotent analysis by presenting the fundamental facts of analysis on the semimodule of mappings of a Hausdorff space into an idempotent semiring A (this semimodule will also be referred to as the space of A-valued functions). In contrast with the original exposition of the theory of idempotent measures and integration [**?**, **?**, **?**], which was modeled on Lebesgue's construction of the measure, here we follow our paper [**?**], which employs ideas close to the construction of Daniell's integral. The latter approach is much easier, since we can immediately use the fact that $\oplus$-summation of infinitely many elements does not cause any trouble in idempotent analysis (in the conventional analysis, it is at this point that one must resort to subtle considerations related to measurability). As a consequence of this simplification, we find that all measures in idempotent analysis are absolutely continuous, that is, can be represented by idempotent integrals with respect to a standard measure with some density function. Actually, this is the main result in this chapter. The other topic discussed here is related to the theory of distributions in idempotent analysis, to the corresponding notions of weak convergence and convergence in measure, and to the relationship between these convergences. The conventional theory of distributions is the main tool in studying solutions of linear equations of mathematical physics; similarly, the theory of new (idempotent) distributions forms a basis for the theory of the Cauchy problem for the differential Bellman equation, which turns out to be linear in the semiring with operations $\oplus = \min$ and $\odot = +$, as well as for quasilinear equations (see Chapter 3).

In radiophysics there is a notion referred to as *signal modulation*, which implies that a lower frequency is imposed on a higher frequency. More precisely, we have a rapidly oscillating function with the upper and the lower envelopes. One of these envelopes (the upper or the lower, depending on the choice of the $\oplus$-operation in the semiring) is the weak limit of the rapidly oscillating function in our sense. For example, the function sequence

$$\varphi(x)\sin nx + \psi(x)\cos nx, \quad \varphi, \psi \in C(\mathbb{R}),$$

is weakly convergent as $n \to \infty$ in the semiring with operations $\oplus = \min$ and $\odot = +$ to the lower envelope $-\sqrt{\varphi^2 + \psi^2}$. Thus, we give a mathematical interpretation of signal modulation.

The Fourier transformation, which provides the eigenfunction expansion of the translation operator, is an important method in the theory of linear equations of mathematical physics. We consider an analog of this transformation in the space of functions ranging in a semiring. Let $A = \mathbb{R} \cup \{+\infty\}$ be the number semiring with operations $\oplus = \min$ and $\odot = +$. Then the eigenfunctions $\psi_\lambda(x)$, $x \in \mathbb{R}^n$, of the translation operator satisfy the equation

$$\psi_\lambda(x+1) = \lambda \odot \psi_\lambda(x) = \lambda + \psi_\lambda(x)$$

and hence are given by the formula $\psi_\lambda(x) = \lambda x$. In the conventional formula for the Fourier transformation, let us replace the usual eigenfunctions $e^{i\lambda x}$ of the translation operator by the idempotent eigenfunctions λx, the multiplication by $\odot = +$, and the integral by the idempotent integral, which is just the operation of taking the infimum. Then we obtain the following analog of the Fourier transformation in our semiring:

$$\widetilde{\varphi}(\lambda) = \inf_x (\lambda x + \varphi(x)).$$

This is the well-known Legendre transformation. The Legendre transformation is a linear operator on the space of functions ranging in a semiring. It has numerous properties resembling those of the usual Fourier transformation; in particular, it takes the $(\oplus, \odot)$-convolution of functions to the multiplication $\odot = +$. The Fourier–Legendre transformation is studied in §1.5. In closing, let us note that idempotent measures (which are the main object in this chapter) provide a rather peculiar example of the general notion of a subadditive (in the usual sense) Choquet capacity.

1.1. Idempotent Semigroups and Idempotent Semirings

Idempotent analysis is based on the notion of an idempotent semiring. In this section we give the definition and provide some examples.

An *idempotent semigroup* is a set M equipped with a commutative, associative operation $\oplus$ (generalized addition) that has a unit element $\mathbb{0}$ such that $\mathbb{0} \oplus a = a$ for each $a \in M$ and satisfies the idempotency condition $a \oplus a = a$ for any $a \in M$. There is a naturally defined partial order on any idempotent semigroup; namely, $a \leq b$ if and only if $a \oplus b = a$. Obviously, the reflexivity of $\leq$ is equivalent to the idempotency of $\oplus$, whereas the transitivity and the antisymmetricity ($a \leq b,\ b \leq a \implies a = b$) follow, respectively, from the associativity and the commutativity of the semigroup operation. The unit element $\mathbb{0}$ is the greatest element; that is, $a \leq \mathbb{0}$ for all $a \in M$. It is also easy to see that the operation $\oplus$ is uniquely determined by the relation $\leq$. Indeed, we have the formula

$$a \oplus b = \inf\{a, b\} \tag{1.1}$$

(recall that the infimum of a subset X in a partially ordered set $(Y, \leq)$ is the element $c \in Y$ such that $c \leq x$ for all $x \in X$ and any element $d \in Y$ satisfying the same condition also satisfies $d \leq c$). Furthermore, if every subset of cardinality 2 in a partially ordered set M has an infimum, then Eq. (1.1) specifies the structure of an idempotent semigroup on M.

Remark 1.1 It is often more convenient to use the opposite order, which is related to the semigroup operation by the formula $a \oplus b = \sup\{a, b\}$.

An idempotent semigroup is called an *idempotent semiring* if it is equipped with yet another associative operation $\odot$ (generalized multiplication) that has a unit element $\mathbb{1}$, distributes over $\oplus$ on the left and on the right, i.e.,

$$a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c), \qquad (b \oplus c) \odot a = (b \odot a) \oplus (c \odot a),$$

and satisfies the property $\mathbb{0} \odot a = \mathbb{0}$ for all a. Distributivity obviously implies the following property of the partial order:

$$\forall a, b, c \quad a \leq b \implies a \odot c \leq b \odot c.$$

An idempotent semiring is said to be *commutative* (*abelian*) if the operation $\odot$ is commutative.

An idempotent semigroup (semiring) M is called an idempotent *metric* semigroup (semiring) if it is endowed with a metric $\rho\colon M \times M \to \mathbb{R}$ such that the following conditions are satisfied: the operation $\oplus$ is (respectively, the operations $\oplus$ and $\odot$ are) uniformly continuous on any order-bounded set in the topology induced by ρ,

$$\rho(a \oplus b, c \oplus d) \leq \max(\rho(a, c), \rho(b, d)) \tag{1.2}$$

(the *minimax axiom*), and

$$a \le b \implies \rho(a,b) = \sup_{c\in[a,b]} \sup(\rho(a,c),\rho(c,b)) \tag{1.3}$$

(the *monotonicity axiom*).

The minimax axiom obviously implies the uniform continuity of $\oplus$ and the minimax inequality

$$\rho\left(\bigoplus_{i=1}^{n} a_i, \bigoplus_{j=1}^{n} b_j\right) \le \min_{\pi} \max_{i} \rho(a_i, b_{\pi(i)}),$$

where the minimum is over all permutations π of the set $\{1,\dots,n\}$. Since the metric is monotone, it follows that any order-bounded set is bounded in the metric.

Let X be a set, and let $M = (M,\oplus,\rho)$ be an idempotent metric semigroup. The set $B(X,M)$ of bounded mappings $X \to M$ (i.e., mappings with order-bounded range) is an idempotent metric semigroup with respect to the pointwise addition $(\varphi\oplus\psi)(x) = \varphi(x)\oplus\psi(x)$, the corresponding partial order, and the uniform metric $\rho(\varphi,\psi) = \sup_x \rho(\varphi(x),\psi(x))$; the validity of axioms (1.2) and (1.3) is obvious. If $A = (A,\oplus,\odot,\rho)$ is a semiring, then $B(X,A)$ bears the structure of an A-*semimodule*; namely, the multiplication by elements of A is defined on $B(X,A)$ by the formula $(a\odot\varphi)(x) = a\odot\varphi(x)$. This A-semimodule will also be referred to as the space of (bounded) A-valued functions on X. If X is a topological space, then by $C(X,A)$ we denote the subsemimodule of continuous functions in $B(X,A)$. If X is finite, $X = \{x_1,\dots,x_n\}$, $n\in\mathbb{N}$, then the semimodules $C(X,A)$ and $B(X,A)$ coincide and can be identified with the semimodule $A^n = \{(a_1,\dots,a_n) : a_j \in A\}$. Any vector $a\in A^n$ can be uniquely represented as a linear combination $a = \bigoplus_{j=1}^{n} a_j \odot e_j$, where $\{e_j,\ j = 1,\dots,n\}$ is the standard basis of A^n (the jth coordinate of e_j is equal to $\mathbb{1}$, and the other coordinates are equal to $\mathbb{0}$). As in the conventional linear algebra, we can readily prove that any homomorphism $m\colon A^n \to A$ (in what follows such homomorphisms are called *linear functionals* on A^n) has the form

$$m(a) = \bigoplus_{i=1}^{n} m^i \odot a_i,$$

where $m^i \in A$. Therefore, the semimodule of linear functionals on A^n is isomorphic to A^n. Similarly, any endomorphism $H\colon A^n \to A^n$ (a linear operator on A^n) has the form

$$(Ha)_j = \bigoplus_{k=1}^{n} h_j^k \odot a_k,$$

i.e., is determined by an A-valued $n \times n$ matrix. By analogy with the case of Euclidean spaces, we define an inner product on A^n by setting

$$\langle a, b\rangle_A = \bigoplus_{k=1}^{n} a_k \odot b_k. \tag{1.4}$$

The inner product (1.4) is bilinear with respect to the operations $\oplus$ and $\odot$, and the standard basis of A^n is orthonormal, that is,

$$\langle e_i, e_j\rangle_A = \delta_A^{ij} = \begin{cases} \mathbb{1}, & i = j, \\ \mathbb{0}, & i \neq j. \end{cases}$$

Let us consider some examples.

Example 1.1 $A = \mathbb{R} \cup \{+\infty\}$ with the operations $\oplus = \max$ and $\odot = +$, the unit elements $\mathbb{0} = +\infty$ and $\mathbb{1} = 0$, the natural order, and the metric

$$\rho(a, b) = |e^{-a} - e^{-b}|.$$

This is the simplest and the most important example of idempotent semiring, involving practically all specific features of idempotent arithmetic. That is why in what follows we often restrict our considerations to spaces of functions ranging in that particular semiring.

Example 1.1′ $A = \mathbb{R} \cup \{-\infty\}$ with the operations $\oplus = \min$ and $\odot = +$. This semiring is obviously isomorphic to that in the preceding example.

Example 1.2 $A = \mathbb{R}_+$ with the operations $\oplus = \min$ and $\odot = \times$ (the usual multiplication). This semiring is also isomorphic to the semiring in Example 1.1; the isomorphism is given by the mapping $x \mapsto \exp(x)$.

Example 1.3 $A = \mathbb{R} \cup \{\pm\infty\}$ with the operations $\oplus = \min$ and $\odot = \max$, the unit elements $\mathbb{0} = +\infty$ and $\mathbb{1} = -\infty$, and the metric

$$\rho(a, b) = |\arctan a - \arctan b|.$$

Example 1.4 Let $\mathbb{R}^n_+$ be the nonnegative octant in $\mathbb{R}^n$ with the inverse Pareto order ($a = (a_1, \dots, a_n) \leq b = (b_1, \dots, b_n)$ if and only if $a_i \geq b_i$ for all $i = 1, \dots, n$). Then $\mathbb{R}^n_+$ is an idempotent semiring with respect to the idempotent addition $\oplus$ corresponding to this order by Eq. (1.1) and the generalized multiplication given by $(a \odot b)_i = a_i + b_i$.

Example 1.5 The subsets of a given set form an idempotent semiring with respect to the operations $\oplus$ of set union and $\odot$ of set intersection. There are various ways to introduce metrics on semirings of sets. For example, if we deal with compact subsets of a metric space, then the Hausdorff metric is appropriate; the distance between two measurable subsets of a space with finite measure can be defined as the measure of their symmetric difference.

Examples 1.4 and 1.5, as well as their various modifications, are used in studying multicriteria optimization problems (see Chapter 3).

Example 1.6 The semiring of endomorphisms of an idempotent semigroup. Let $M = (M, \oplus, \rho)$ be an idempotent metric semigroup. Let $E = \mathrm{End}(M)$ be the set of endomorphisms of M, i.e., continuous mappings $h: M \to M$ such that $h(a \oplus b) = h(a) \oplus h(b)$ for any $a, b \in M$. We define the pointwise generalized addition $(h \oplus g)(a) = h(a) \oplus g(a)$ on E in the usual way; the multiplication on E is defined as the composition of mappings, i.e., $(h \odot g)(a) = h(g(a))$. The unit elements for these operations are the absorbing endomorphism $\mathbb{0}_E$ given by $\mathbb{0}_E(a) = \mathbb{0}$ for any a and the identify endomorphism $\mathbb{1}_E$ given by $\mathbb{1}_E(a) = a$. The set E equipped with these operations is an idempotent semiring. For example, if $M = \mathbb{R} \cup \{+\infty\}$ is the number semigroup with operation $\oplus = \min$ and with the metric discussed in Example 1.1, then E is just the set of monotone continuous self-mappings of $\mathbb{R} \cup \{+\infty\}$. Next, let M satisfy the additional condition that all balls $B_R = \{a \in M : \rho(a, \mathbb{0}) \le R\}$ are compact (as is the case for $M = \mathbb{R} \cup \{+\infty\}$). Then the metric

$$\rho(h, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\rho_n(h, g)}{1 + \rho_n(h, g)},$$

where $\rho_n(h, g) = \sup\{\rho(h(a), g(a)) : a \in B_n\}$, is defined on $E = \mathrm{End}(M)$. The axioms of metric and the minimax and the monotonicity conditions (1.2) and (1.3) follow by straightforward verification, and we see that $\mathrm{End}(M)$ is an idempotent metric semiring.

Example 1.7 The semiring of endomorphisms of (or linear operators on) a function semimodule. A linear operator on the space $B(X, A)$ of bounded A-valued functions (or on the subspace $C(X, A) \subset B(X, A)$ if X is a topological space) is a continuous mapping $H: B(X, A) \to B(X, A)$ (respectively, $H: C(X, A) \to C(X, A)$) such that

$$H(a \odot h \oplus b \odot g) = a \odot H(h) \oplus b \odot H(g)$$

for any functions h and g and any constants a and b. The set of such operators is an idempotent metric semiring with respect to the operations $\oplus$ of pointwise addition and $\odot$ of composition (as in Example 1.6) and with the distance between two operators determined by the formula

$$\rho(H_1, H_2) = \sup\{\rho(H_1(f), H_2(f)) : f(x) \in [\mathbb{1}, \mathbb{0}] \ \forall x\}.$$

The usual multiplication by elements of A turns the semiring of operators into a semialgebra, which is the idempotent analog of the von Neumann algebra. As was indicated above, for a finite $X = \{x_1, \dots, x_n\}$ the function space is A^n and the operator semialgebra is isomorphic to the semialgebra of A-valued $n \times n$ matrices. For these matrices, metric convergence is equivalent to strong convergence (which readily follows from the fact that the metric on A satisfies the minimax axiom and from the uniform continuity of $\oplus$ and $\odot$); in turn, strong convergence is equivalent to coordinatewise convergence. Thus, for any matrices B_k and B we have

$$\begin{aligned} B_k \to B \text{ as } k \to \infty \iff & (B_k)_{ij} \to B_{ij} \text{ in } A \text{ for any } i, j \\ \iff & B_k h \to Bh \text{ in } A^n \text{ for any } h \in A^n. \end{aligned}$$

Example 1.8 Convolution semirings. If X is a topological group and A is an idempotent semiring in which any bounded subset has an infimum, then we can define an idempotent analog $\circledast$ of convolution on $B(X, A)$ by setting

$$(\varphi \circledast \psi)(x) = \inf_{y \in X} (\varphi(y) \odot \psi(x - y)). \tag{1.5}$$

This operation turns $B(X, A)$ into an idempotent semiring, which will be referred to as the *convolution semiring*.

1.2. Idempotent Linear Algebra, Graph Optimization Algorithms, and Discrete Event Dynamical Systems

This section contains a brief exposition of idempotent linear algebra (the theory of finite-dimensional semimodules over semirings with idempotent addition) and its applications to the construction and analysis of discrete optimization problems on graphs. A comprehensive discussion of the topic, including a wide variety of specific practical problems, algorithms, and methods, can be found in [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**].

Throughout this section X is a finite set of cardinality n and A is an idempotent metric semiring.

Let H be a linear operator with matrix (h^i_j) on $C(X, A)$, and let $F \in C(X, A)$. The discrete-time equation

$$S_{t+1} = HS_t \oplus F, \qquad S_t \in C(X, A),\ t = 0, 1, \dots, \tag{1.6}$$

in the function space $C(X, A)$ is called the *generalized evolution Bellman equation*, and the equation

$$S = HS \oplus F \tag{1.7}$$

is called the *generalized stationary* (*steady-state*) *Bellman equation*. These equations are said to be *homogeneous* if $F = \mathbb{0}$ and *nonhomogeneous* otherwise. In other words, generalized discrete Bellman equations are general linear equations in A^n. To each Bellman equation there corresponds a geometric object called a *discrete medium* or a *graph*; this object can be described as follows. The elements of the set $X = \{x_1, \dots, x_n\}$ are referred to as the *points* of the medium; each ordered pair $(x_i, x_j) \in X \times X$ such that $h^i_j \neq \mathbb{0}$ is called a *link* between x_i and x_j. Let $\Gamma \subset X \times X$ be the set of all links. The mapping $L\colon \Gamma \to A \setminus \{\mathbb{0}\}$ given by the formula $L(x_i, x_j) = h^i_j$ is called the *link characteristic*, and the discrete medium is the quadruple $M = (X, \Gamma, L, A)$. In alternative terms, X is the set of nodes, Γ is the set of (directed) arcs, and h^i_j are the arc weights; $M(X, \Gamma, L, A)$ is a weighted directed graph with arc weights in A. Set $\Gamma(x_i) = \{x' \in X : (x', x_i) \in \Gamma\}$. Then $\Gamma(x_i)$ is the set of points linked with x_i; it will be called the *neighborhood* of x_i. The operator H can be represented as

$$(H\varphi)(x_i) = \bigoplus_{x' \in \Gamma(x_i)} L(x', x_i) \odot \varphi(x'). \tag{1.8}$$

This is just the form in which H and the corresponding equations (1.6) and (1.7) arise when optimization problems on graphs are solved by dynamic programming. Typically, the semiring A in these applications is one of the semirings described in Examples 1.1–1.3.

Equations that are only additive (but not linear) also occur in practical discrete optimization problems very frequently. They can be viewed as a straightforward generalization of Eqs. (1.6) and (1.7). Let us give the precise definition. Suppose that M is an idempotent metric semigroup, $F = \{F_j\}$ is a function in $C(X, M) = M^n$, and $H \in \operatorname{End}(M^n)$ is an arbitrary endomorphism (see Example 1.6 above). Then Eq. (1.7) for the unknown function $S \in C(X, M)$ is called the stationary Bellman equation in the space of functions ranging in the idempotent semigroup M. The evolution equation is defined similarly. To write out this equation more explicitly, we use the following obvious statement.

Proposition 1.1 *The semiring* $\operatorname{End}(M^n)$ *of endomorphisms of the semigroup* M^n *is isomorphic to the semiring of linear operators on the semimodule* $(\operatorname{End}(M))^n$*, i.e., to the semiring of* $\operatorname{End}(M)$*-valued* $n \times n$ *matrices, so that a general element* $H \in \operatorname{End}(M^n)$ *is given by the formula*

$$(H\varphi)_i = \bigoplus_{j=1}^{n} H_{ij}(\varphi_j), \qquad \varphi = (\varphi_1, \dots, \varphi_n) \in M^n,$$

where all H_{ij} *belong to* $\operatorname{End}(M)$.

A generalization of this assertion to function semigroups $C(X, M)$ with infinite X is given in §1.3. Proposition 1.1 obviously implies that the general stationary Bellman equation on the space of functions with values in an idempotent semigroup has the form

$$\varphi_i = \bigoplus_{j=1}^{n} H_{ij}(\varphi_j) \oplus F_i,$$

or

$$\varphi(x_i) = \bigoplus_{j=1}^{n} H_{ij}(\varphi(x_j)) \oplus F(x_i). \tag{1.9}$$

Note that the Bellman equation on the space of functions ranging in a semiring is a special case of Eq. (1.9), in which M is a semiring and the endomorphisms H_{ij} are left translations, i.e., $H_{ij}(\varphi_j) = h_j^i \odot \varphi_j$ for some $h_j^i \in M$.

Associated with Eq. (1.9) is the operator equation

$$\widehat{\varphi} = H\widehat{\varphi} \oplus \widehat{F} \tag{1.10}$$

for an unknown operator-valued function $\widehat{\varphi} \in C(X, \mathrm{End}(M)) = (\mathrm{End}(M))^n$, where $\widehat{F} \in C(X, \mathrm{End}(M))$ is given. Although the properties of Eqs. (1.9) and (1.10) are quite similar (e.g., see Proposition 1.2 below), Eq. (1.10) is linear over the semiring (note that Eq. (1.9) is only additive), which simplifies its study dramatically. Hence, we can say that Proposition 1.1 permits us to reduce additive Bellman equations for semigroup-valued functions to linear Bellman equations for semiring-valued functions.

By analogy with the preceding, we can interpret Eq. (1.9) geometrically as an equation in a discrete medium. Namely, the link characteristic $L(x_i, x_j) = H_{ij}$ ranges over the semiring of endomorphisms, and so Eq. (1.9) can be represented in the form

$$S(x_j) = (HS)(x_j) \oplus F(x_j), \tag{1.9$'$}$$

where the endomorphism $H\colon C(X, M) \to C(X, M)$ is determined by the equation

$$(HS)(x_i) = \bigoplus_{x' \in \Gamma(x_i)} L(x', x_i) S(x') \tag{1.11}$$

(here $\Gamma(x_i) = \{x' : L(x', x_i) \neq \mathbb{0}_E\}$ is the neighborhood of the point x_i of the medium).

Let us now derive the basic formulas for the solutions of the Bellman equation directly from linearity or additivity.

First, the solution S_t^F of the evolution equation (1.6) with the zero initial function $S_0 = \mathbb{0}$ obviously has the form

$$S_t^F = H^{(t-1)} F \equiv \bigoplus_{k=0}^{t-1} H^k F \tag{1.12}$$

in both the linear and the additive case (here H^k is the kth power of H). This is the simplest Duhamel type formula expressing a particular solution of the nonhomogeneous equation as a generalized sum of solutions of the homogeneous equation.

Now, if the equation is A-linear, then the solution $S_t = H^t S_0$ of the homogeneous equation (1.6) with $F \equiv \mathbb{0}$ and with an arbitrary initial function $S_0 = \bigoplus_{i=1}^n S_0^i \odot e_i$, where e_i is the standard basis in A^n (see 1.1), is obviously equal to

$$S_t = \bigoplus_{i=0}^{n} S_0^i \odot H^t e_i,$$

that is, is a linear combination of the source functions (Green's functions) $H^t e_i$, which are the solutions of Eq. (1.6) with the localized initial perturbations $S_0 = e_i$. Similarly, if $S = G_i \in A^n$ is the solution of the stationary equation (1.7) with localized nonhomogeneous term $F = e_i$, i.e., $G_i = HG_i \oplus e_i$, then by multiplying these equations by arbitrary coefficients $F_i \in A$ and by taking the $\oplus$-sum of the resultant products, we find that the function

$$S = \bigoplus_{i=1}^{n} F_i \odot G_i$$

is a solution of Eq. (1.7) with arbitrary $F = \bigoplus F^i \odot e_i$. Thus, we have obtained a source representation of solutions of the stationary A-linear Bellman equation.

Let us return to the general additive equation (1.9) and to the corresponding equation (1.10). The operator $H = (H_{ij})$ occurring in these equations is linear in $(\mathrm{End}(M))^n$. Recall that in the linear equation (1.7) H is determined by a matrix h^i_j of elements of the semiring A. A majority of methods for constructing solutions of the stationary Bellman equation are based on the following simple assertion.

Proposition 1.2 *Set* $H^{(t)} = \bigoplus_{k=0}^{t} H^k = (H \oplus I)^t$, *where* I *is the identity operator. If the sequence* $H^{(t)}$ *converges to a limit* H^* *as* $t \to \infty$, $t \in \mathbb{N}$ (*the convergence of operators is defined in Example* 1.6), *then* $\varphi = H^* F$ *and* $\widehat{\varphi} = H^* \widehat{F}$ *are solutions of Eqs.* (1.9) *and* (1.10), *respectively.*

The solution $S^F_* = H^* F$ of Eq. (1.9) is called the *Duhamel solution*.

The proof is by straightforward substitution. For example, for Eq. (1.9) we have

$$\begin{aligned} H(S^F_*) \oplus F &= H \lim_{t\to\infty} (H^{(t)} F \oplus F) \\ &= \lim_{t\to\infty} H((H^t F) \oplus F) = \lim_{t\to\infty} H^{(t+1)} F = S^F_*. \end{aligned}$$

An important supplement to Proposition 1.2 is given by the following assertion.

Proposition 1.3 *Let the operator sequences* $H^{(t)}$ *and* H^t *tend to some limits* H^* *and* H^∞, *respectively, as* $t \to \infty$. *Then each solution of Eq.* (1.9) *has the form*

$$\varphi = g \oplus S^F_*, \tag{1.13}$$

where g *is a solution of the homogeneous equation* $Hg = g$. *Moreover,* S^F_* *is a unit element with respect to the operation* $\oplus$ *on the set of all solutions of Eq.* (1.9), *i.e., for each solution* S *one has* $S = S \oplus S^F_*$.

Proof. Let φ be a solution of Eq. (1.9). On substituting $\varphi = H\varphi \oplus F$ into the right-hand side, we obtain

$$\varphi = H^2\varphi \oplus HF \oplus F.$$

After k successive applications of this procedure, we have

$$\varphi = H^k\varphi \oplus H^{(k-1)}F.$$

This is valid for all $k \in \mathbb{N}$. By passing to the limit as $k \to \infty$, we obtain

$$\varphi = H^\infty\varphi \oplus S_*^F,$$

which proves Eq. (1.13), since $H^\infty\varphi$ is obviously a solution of the homogeneous equation $H(H^\infty\varphi) = H^\infty\varphi$. The second part of the proposition readily follows from Eq. (1.13) and from the fact that the operation $\oplus$ is idempotent.

Thus, to solve the stationary equation (1.9) effectively, one should be capable of evaluating the limits of the operator sequences $H^{(t)} = (H \oplus I)^t$ and H^t. Hence, it is important to have some criteria for the existence of these limits and for the possibility to evaluate them in a finite number of iterations (the latter is needed if computers are to be used and computational complexity is to be estimated). These criteria will be discussed later. Meanwhile, let us give another important representation of solutions of the evolution equation; this representation will be used in the subsequent discussion.

If we use the standard rules of matrix multiplication to write out the powers of matrices in (1.12) explicitly, then we obtain

$$(S_t^F)_i = \bigoplus_{k=0}^{t-1} \bigoplus_{j_1,\dots,j_k} H_{ij_1}H_{j_1j_2}\dots H_{j_{k-1}j_k}F_{j_k}, \quad i, j_1, \dots, j_k = 1, \dots, n. \tag{1.14}$$

Let us now rewrite this formula in terms of the discrete medium on the basis of representation (1.8) or (1.11) of the operator H. To this end, we introduce some notation. A *path* $\mu_{x,y}^k = \{x_0, x_1, \dots, x_k\}$ *of length* $|\mu| = k$ joining two points x and y of the medium X is a sequence of $k+1$ points of X such that $x_0 = x$, $x_k = y$, and for each pair (x_i, x_{i+1}) of neighboring points we have $(x_i, x_{i+1}) \in \Gamma$ (that is, each such pair is a link, or an arc of the graph). Let $M_{x\to y}^k$ (respectively, $M_{x\to y}^{(k)}$ or $M_{x\to y}$) denote the set of all paths of length k (respectively, of length $\le k$ or of arbitrary length) joining the points x and y. We define the *path endomorphism*

$$L(\mu_{x\to y}^k) \colon C(X, M) \to C(X, M)$$

corresponding to a path $\mu_{x\to y}^k$ by setting

$$L(\mu_{x\to y}^k) = L(x_k, x_{k-1})L(x_{k-1}, x_{k-2})\dots L(x_1, x_0).$$

We write $\operatorname{supp}_{\mathbb{0}} F = \{x_i : F(x_i) \neq \mathbb{0}\}$; this set is called the support of F. Obviously, Eq. (1.13) for S_t^F can be rewritten in the form

$$(S_t^F)(x) = \bigoplus_{y \in \operatorname{supp}_{\mathbb{0}} F} \bigoplus_{\mu \in M_{y \to x}^{(t)}} L(\mu) F(y). \tag{1.15}$$

If the sequence S_t^F converges as $t \to \infty$ to some limit S_*^F, then we have the similar representation

$$(S_*^F)(x) = \bigoplus_{y \in \operatorname{supp}_{\mathbb{0}} F} \bigoplus_{\mu \in M_{y \to x}} L(\mu) F(y). \tag{1.16}$$

Representations (1.15) and (1.16) of solutions of linear or additive Bellman equations by $\oplus$-sums of contributions of all paths are idempotent discrete counterparts of the path integral representation of solutions of linear differential equations.

According to Proposition 1.2, to find the Duhamel solution S_*^F of the stationary Bellman equation (either linear or additive), one has to calculate the limit of the operator sequence $H^{(t)} = \bigoplus_{k=0}^t H^k = (H \oplus I)^k$, where H is a linear operator on the semimodule $C(X, A)$; moreover, if the Bellman equation under consideration is only additive, then the semiring A should be chosen as the semiring of endomorphisms of the corresponding idempotent semigroup. For the limit of $H^{(t)}$ to be computable in finitely many steps, we require the sequence $H^{(t)}$ to stabilize starting from some term.

Definition 1.1 We say that an operator $H: C(X, A) \to C(X, A)$ is *nilpotent* if there exists a $t_0 \in \mathbb{N}$ such that $H^{(t_0)} = H \oplus H^{(t_0)} \oplus \mathbb{1}$, where $\mathbb{1}$ is the identity operator.

Obviously, if H is nilpotent, then the sequence $H^{(t)}$ stabilizes, i.e., $H^{(t)} = H^{(t_0)}$ for all $t \geq t_0$, and consequently, the Duhamel solution exists and can be found in finitely many operations. Let us give a nilpotency criterion in terms of the geometry of the discrete medium (X, L, Γ, A) corresponding to the operator H.

Let M_c be the set of elementary circuits in the graph (X, Γ), i.e., nonself-intersecting closed paths in the medium.

The following result is well known; e.g., see [**?**, **?**, **?**, **?**].

Theorem 1.1 *If*

$$\mathbb{1} \oplus L(\nu) = \mathbb{1} \tag{1.17}$$

for each circuit $\nu = (x_0, x_1, \dots, x_k = x_0) \in M_c$*, where* $L(\nu)$ *is the path endomorphism and* $\mathbb{1}$ *is the identity operator in* $C(X, A)$*, then the operator* H *is nilpotent.*

The proof follows from representation (1.15) of the solution $S_t^F = H^{(t)} F$. Indeed, by virtue of (1.17) the value of the right-hand side of (1.15) remains the same if in the $\oplus$-sum we retain only the contributions of nonself-intersecting paths, i.e., paths that do not contain elementary circuits. Obviously, there are finitely many such paths. Hence, there exists a t_0 such that $H^{(t)}F = H^{(t_0)}F$ for all $t > t_0$ and all F, which completes the proof.

The theorem gives a sufficient condition for H to be nilpotent. However, if A satisfies the additional property that it does not have the least element and for any $a \in A$ with $a \oplus \mathbb{1} \neq \mathbb{1}$ and any $b \in A$ there exists an $n \in \mathbb{N}$ such that $a^n = a \odot \cdots \odot a \leq b$ (for example, this property is valid for the number semiring in Example 1.1), then the condition given in the theorem is also necessary for the operator sequence $H^{(k)}$ to converge as $k \to \infty$.

Another type of nilpotency criteria is provided by spectral analysis. We say that a vector $v = (v_i) \in A^n = C(X, A)$ and an element $\lambda \in A$ are an eigenvector and the corresponding eigenvalue of H, respectively, if

$$Hv = \lambda \odot v. \tag{1.18}$$

As in the usual linear algebra, there exists a close relationship between the behavior of the iterations of H and its eigenvalues and eigenvectors. However, the study of Eq. (1.18) in the framework of idempotent analysis reveals some effects that do not occur in the usual algebra and functional analysis. Let us state a simple result for the case of the number semiring $A = \mathbb{R} \cup \{+\infty\}$ from Example 1.1; this result elucidates the characteristic features of idempotent spectral analysis.

The following theorem is well known (e.g., see [**?**, **?**, **?**]).

Theorem 1.2 *For any linear operator $H: A^n \to A^n$, where $A = \mathbb{R} \cup \{+\infty\}$, there exists a $\lambda \in A$ and a $v \in A^n$, $v \neq \mathbb{0}$, such that Eq.* (1.18) *is satisfied. Furthermore, if all matrix elements h^i_j of H satisfy $h^i_j \neq \mathbb{0}$, then $\lambda \neq \mathbb{0}$ and is unique, and the vector v is unique up to $\odot$-multiplication by a constant. The operator H is nilpotent if $\lambda \geq \mathbb{1} = 0$ and is not nilpotent if $\lambda < \mathbb{1}$.*

The proof of this theorem, as well as a generalization to the function space $C(X, A)$ over an infinite set X, will be given in §2.3, where the spectral properties of idempotent linear operators are discussed in detail.

If H is nilpotent, then, as was shown above, we can obtain a finite algorithm for constructing the Duhamel solution S^F_* of the stationary Bellman equation by successively evaluating the powers of H. This method for constructing S^F_* is known as the Picard–Bellman method. However, it is often more efficient to find S^F_* by using a generalization of the Ford–Lee wave algorithm to an arbitrary semiring. Let us briefly present this algorithm as well as some typical examples of problems that can be solved by this method. We essentially follow [**?**, **?**] (see also [**?**]).

To find the Duhamel solution S^F_* of Eq. (1.9), we introduce additive operators $\pi_{ij}: C(X, M) \to C(X, M)$ by setting

$$(\pi_{ij}\varphi)(x) = \begin{cases} \varphi(x), & x \neq x_j, \\ H_{ij}(\varphi(x_i)) \oplus \varphi(x_j), & x = x_j. \end{cases}$$

These operators will be referred to as the *local modifications* along the links $(x_i, x_j) \in \Gamma$. We define a sequence $\{S_k\}$, $k = 0, 1, \ldots$, of functions in $C(X, M)$ by the following recursion law: we set $S_0 = F$ and

$$S_t = \begin{cases} S_{t-1} & \text{if the set } Q_t = \{(i,j) : S_{t-1} \neq \pi_{ij} S_{t-1}\} \text{ is empty,} \\ \pi_{ij} S_{t-1} & \text{if } Q_t \neq \emptyset \end{cases}$$

(here (i, j) is an arbitrary element of Q_t). It turns out that under the nilpotency condition (1.17) this sequence (which is determined ambiguously since the choice of $(i, j) \in Q_t$ is not specified) stabilizes; that is, there exists a t_0 such that $S_{t_0} = S_*^F$. The proof of this statement is essentially similar to the proof of the fact that the iterations H^k stabilize.

Concrete algorithms implementing the cited scheme can be obtained by specifying the method of choosing the link $(i, j) \in Q_t$.

In all Bellman equations that correspond to the known classical optimization problems on graphs, the operation $\oplus$ is induced by a linear ordering relation and hence satisfies the extended idempotency law: the $\oplus$-sum of any two (and thus of any finite number of) elements of the semigroup is equal to the smaller element. Hence, $S_*^F(x)$ in Eq. (1.16) coincides with $L(\mu_0)F(y_0)$ for some path $\mu_0 \in M_{y_0 \to x}$, referred to as the *optimal* path. In this situation, the modifications π_{ij} of the medium states can be accompanied by path modifications Π_{ij}, just as in the conventional Ford–Lee wave method. Namely, with the sequence S_t of states (i.e., of functions in $C(X, M)$) we associate a sequence ω_t of functions from $X \setminus \operatorname{supp}_{\mathbb{0}} F$ to X as follows. Each modification π_{ij} of the state S_{t-1} induces a modification Π_{ij} of the function ω_{t-1} according to the formula

$$\omega_t(x) = \Pi_{ij}\omega_{t-1}(x) = \begin{cases} \omega_{t-1}(x), & x \neq x_j, \\ x_i, & x = x_j; \end{cases}$$

the choice of the initial function ω_0 can be arbitrary. Thus, the extended algorithm finds the solution S_*^F (the Bellman function) as well as the optimal path (on which the infimum in (1.16) is attained). Indeed, it is almost obvious that if $S_*^F = S_t$, then the optimal path from the set $\Phi_0 = \operatorname{supp}_{\mathbb{0}} F$ to any point $x \in X \setminus \Phi_0$ can be reconstructed from the function ω_t recursively, according to the following rule: $(\omega_t(x), x)$ is the last arc in this path.

Let us consider some typical examples.

1. A classical trajectory-type problem. In this problem we consider a medium (X, Γ, L, A) with a semiring A in which the operation $\oplus$ is induced by a linear ordering relation $\leq$. Suppose that a function $F \in C(X, A)$ with support

$$\operatorname{supp}_{\mathbb{0}} F = \Phi_0 = \{x : F(x) \neq \mathbb{0}\}$$

and a set $\Phi \subset X$ such that $\Phi \cap \Phi_0 = \emptyset$ are given. Let $M_{\Phi_0 \to \Phi}$ be the set of paths joining Φ_0 and Φ, that is,

$$M_{\Phi_0 \to \Phi} = \bigcup_{x \in \Phi_0} \bigcup_{y \in \Phi} M_{x \to y}.$$

The problem is to find an optimal path $\mu_0 \in M_{\Phi_0\to\Phi}$ with initial point $x_0 \in \Phi_0$ and final point $y \in \Phi$; the optimality is understood in the sense that

$$L(\mu_0)F(x_0) \le L(\mu)F(x_0) \quad \forall\, \mu \in M_{\Phi_0\to\Phi}.$$

This problem can be reduced to finding the Duhamel solution S_*^F of the Bellman equation (1.7) (which is linear in the semiring A) with nonhomogeneous term F and a path μ_0 for which $S_*^F(x) = L(\mu_0)F(x_0)$; the latter problem can be solved by the generalized wave method. Various choices of the semiring A correspond to various concrete statements of the problem. Practically, problems corresponding to the semirings described in Examples 1.1–1.3 are considered most often. However, other semirings also sometimes occur in such problems; e.g., let us mention the semiring $A = \{0, 1, \dots, m\}$ with operations $a \oplus b = \min(a, b)$ and $a \odot b = \min(m, a + b)$.

2. A nonclassical trajectory-type problem (the shortest path in a network with variable arc traversing time). Suppose that to each arc $u \in \Gamma$ of a graph $G(X, \Gamma)$ there corresponds a nonnegative function $\varphi_u(t)$ specifying the time in which u is traversed if the motion begins at time t. The problem is to construct a time-optimal path $\mu_0(t_0)$ from a marked node $x_0 \in X$ to any other node $x \in X$ under the condition that the motion from x_0 begins at time t_0. The corresponding discrete medium is determined by the graph (X, Γ), the semigroup $M = \{t \in \mathbb{R} : t \ge t_0\}$ with operation $\oplus = \min$, and the endomorphisms $L(u) \in \mathrm{End}(M)$ given by the formula

$$(L(u))(a) = \min_{\tau \ge a}\{\tau + \varphi_u(\tau)\}.$$

It is easy to show that the length S_*^F of the shortest path from $x_0 \in X$ to $x \in X$ is a solution of the Bellman equation (1.9′) with the right-hand side

$$F(x) = \begin{cases} t_0, & x = x_0, \\ \mathbb{0} = +\infty, & x \ne x_0. \end{cases}$$

It turns out that the generalized wave method offers a more efficient solution of this problem than the previously used dynamic programming algorithm with discrete time.

3. The problem of product classification by construction-technological criteria. One usually solves this problem by methods of cluster analysis. These methods actually use concrete semirings determined by the specific methods of defining the distance between objects in the criterion space. The problem is stated as follows: find a partition $X = X_1 \cup \dots \cup X_m$, $X_i \cap X_j = \emptyset$, $X_j \ne \emptyset$, of a finite set X of cardinality n into $m \le n$ nonempty clusters so as to minimize the sum $\sum_{i=1}^m D(X_i)$ of generalized distances between the objects within each cluster. Here $D(X_i) = \sum_{x,y\in X_i} d(x, y)$, where $d\colon X \times X \to \mathbb{R}_+$ is a given function. This problem is linear in the space of mappings of the set of pairs

$$(Y, j), \quad Y \subset X, \quad j \in \mathbb{N}, \quad 1 \le j \le |Y|, \quad |Y| \ne \emptyset,$$

into the semiring $(\mathbb{R}_+, \oplus = \min, \odot = +)$, and the solution is determined by the solution of the Bellman equation

$$S(Y, j) = \min\Big(\min_{\emptyset \neq Z \subset Y}(S(Y \setminus Z, j-1) + D(Z)), F(Y, j)\Big). \tag{1.19}$$

Here $S(Y, j)$ is the sum of generalized distances for the optimal partition of Y into j subsets and

$$F(Y, j) = \begin{cases} D(X), & j = 1, \\ \mathbb{0} = +\infty, & j \neq 1. \end{cases}$$

In concrete classification problems, the generalized metric is determined by the numerical values of the parameters that specify the elements of X in the criterion space. The classification problem remains linear even if the metric d is induced by odd binary relations between objects.

4. The generalized assignment problem: a multi-iteration algorithm. It is sometimes convenient to solve optimization problems on graphs by constructing solutions of several Bellman equations in media varying from step to step. By way of example, let us consider the generalized assignment problem for a complete bipartite graph. Let $G = (X, \Gamma)$ be a complete bipartite graph; i.e., $X = Y \cup Z$, $Y \cap Z = \emptyset$, $|Y| = |Z| = n$, $\Gamma = Y \times Z$. A *matching* Π is an arbitrary set of pairwise nonadjacent arcs. Let A be a semiring with operation $\oplus$ induced by a linear ordering relation $\leq$, and let a *weight function* $f\colon \Gamma \to A$ (or a *link characteristic function*) be given. We define the *weight* of a matching Π by setting $f(\Pi) = \odot_{e\in\Pi} f(e)$. A matching Π_k^0 of cardinality k is said to be *optimal* if $f(\Pi_k^0) \leq f(\Pi_k)$ for any other matching Π_k of cardinality k. The problem is to find an optimal matching of maximal cardinality.

Assume in addition that the operation $\odot$ is either a group operation on A (i.e., inverse elements always exist) or an idempotent operation. The multi-iteration algorithm successively constructs optimal matchings of cardinality $k = 1, \dots, n$, so that the number of iterations is equal to n. On the first step we choose an arc $e_1 = (y_1, z_1) \in \Gamma$ of minimal weight and reverse it, i.e., replace e_1 by the arc $\tilde{e}_1 = (z_1, y_1)$; the weight assigned to $\tilde{e}_1$ is $f(e_1)^{-1}$ (if $\odot$ is a group operation) or $f(e)$ (if $\odot$ is an idempotent operation). Thus, Γ is transformed into another graph G_1 (which is obviously not bipartite). This completes the first iteration. On the second step we seek a minimal-weight path from $Y \setminus \{y_1\}$ to $Z \setminus \{z_1\}$ in G_1. To this end, we solve the corresponding Bellman equation from Example 1.1 by the generalized wave algorithm. Let y_2 be the first node and z_2 the last node of this path. We reverse all arcs in this path, thus obtaining a new graph G_2 and completing the second iteration. On the next step we seek a minimal-weight path from $Y \setminus \{y_1, y_2\}$ to $Z \setminus \{z_1, z_2\}$ in G_2, and so on. By induction on k, using the alternating path method, we can show that the set of reversed arcs forms an optimal matching of cardinality k on the kth iteration.

Let us give a sketch of proof. Let Π_k and Π_{k+1} be matchings of cardinality k and $k+1$, respectively. An *alternating path* is a path in the nondirected version of G such that the arcs in this path alternately lie in Π_k and Π_{k+1} and the first node, as well as the last node, is incident to exactly one arc in the union $\Pi_k \cup \Pi_{k+1}$. In these terms, our statement is equivalent to the following: there exists an optimal matching $\widetilde{\Pi}_{k+1}$ such that the set

$$\Pi_k \triangle \widetilde{\Pi}_{k+1} = (\widetilde{\Pi}_{k+1} \setminus \Pi_k) \cup (\Pi_k \setminus \widetilde{\Pi}_{k+1})$$

is an alternating path. But it is clear that for any matchings Π_k and Π_{k+1} the set $\Pi_k \cup \Pi_{k+1}$ is a disjoint union of alternating paths. Thus, to prove the desired assertion, we can start from arbitrary optimal matchings Π_k and Π_{k+1} and successively modify Π_{k+1} so that the number of alternating paths in $\Pi_k \triangle \Pi_{k+1}$ is reduced to one. This is quite easy. For example, let C be a path with even number of arcs in $\Pi_k \triangle \Pi_{k+1}$. Obviously, the $\odot$-product of the weights of arcs in $C \cap \Pi_k$ is equal to the $\odot$-product of the weights of arcs in $C \cap \Pi_{k+1}$, and we can replace Π_{k+1} by the matching

$$\Pi'_{k+1} = (\Pi_{k+1} \setminus C) \cup (C \cap \Pi_k),$$

which is also optimal.

Typically, the semirings from Examples 1.1–1.3 are used in this problem.

Similarly, one can construct multi-iteration algorithms for flow and transport problems and for various optimal performance problems related to flexible manufacturing systems (e.g., see [**?**]).

5. The generalized assignment problem. A single-iteration algorithm. Numerous problems traditionally solved by multi-iteration algorithms can be solved by single-iteration algorithms under an appropriate choice of the function semimodule. As an example, let us consider the cited generalized assignment problem. Let $Y_k \subset Y$ and $Z_k \subset Z$ be subsets of cardinality k, $k = 1, \dots, n$, and let $S(k, Y_k, Z_k)$ be the weight of an optimal matching of cardinality k in the subgraph of G generated by $Y_k \cup Z_k$. Then S obviously satisfies the following Bellman equation, linear in the semiring A:

$$S(k, Y_k, Z_k) = \bigoplus_{\substack{y \in Y_k, \\ z \in Z_k}} S(k-1, Y_k \setminus \{y\}, Z_k \setminus \{z\}) \odot f(y,z) \oplus F(k, Y_k, Z_k),$$

where $F(1, y, z) = f(y, z)$ and $F(k, Y_k, Z_k) = \mathbb{0}$ for $k > 1$.

6. The modified traveling salesman problem in a computational medium. In a computational medium $G = (X, \Gamma, L, A)$ with arc weights in a semiring A, it is required to find the shortest path passing just once through each node of X (the beginning and the end of the path are not specified; in contrast to the usual traveling salesman problem, the path need not be closed). Let (Y, y) be a pair of a subset $Y \subset X$ and a distinguished node

$y \in Y$, and let $S(Y, y)$ be the weight of the shortest path issuing from Y, lying in Y, and visiting each node in Y exactly once. Then S satisfies the Bellman equation

$$S(Y, y) = \bigoplus_{z \in Y \setminus \{y\}} L(y, z) \odot S(Y \setminus \{y\}, z) \oplus F(Y, y),$$

where $F(Y, y) = \mathbb{1}$ if $Y = \{y\}$ is a singleton and $F(Y, y) = \mathbb{0}$ otherwise.

We point out that this problem is NP-hard. However, there exists a heuristic polynomial algorithm for this problem with accuracy estimate depending on the admissible solution time [**?**].

Problems corresponding to this mathematical formulation very frequently occur in technics. For example, we can indicate the minimal time equipment resetting problem (for a given set $X = \{x_1, \dots, x_n\}$ of tasks and given times t_{ij} of equipment resetting for x_j to be performed after x_i, it is required to find the order of tasks that minimizes the total resetting time), the problem of determining the optimal order of assembling printed circuit-based units, the problem of optimal motion of a piler robot, etc.

7. Analysis of parallel computations and production scheduling. The mathematical aspects of the theory of computational media are studied in the book [**?**], where specific parallel computation algorithms are constructed and thoroughly analyzed (synthesis of optimal data loaders for matrix processors, a parallel LU-decomposition algorithm, etc.). We restrict ourselves to the simplest examples of mathematical problems that arise here (for detail, see [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**] and references therein).

Let us schematically describe a solution algorithm for some problem by the following data: a set $X = \{x_1, \dots, x_n\}$ of elementary operations involved in the algorithm, a strict partial order on X ($x_i < x_j$ implies that the operation x_j cannot be initiated before the operation x_i is finished), and a function $T: X \to \mathbb{R}_+$ that determines the execution time for each operation. We assume that data transfers between the processors are instantaneous and that each operation x_j can be executed by any processor in the same time $T(x_j)$.

First, suppose that the computational medium contains enough processors to load the programs corresponding to all operations simultaneously. Then each operation is initiated immediately after all preceding operations have been finished. Let t_0 be the starting time of the algorithm, and let $S(x_j)$ be the termination time of the operation x_j. Obviously, the function $S: X \to \mathbb{R}$ is a solution of the nonhomogeneous discrete Bellman equation

$$S(x_j) = \max\Big(\max_{x_i < x_j} S(x_i) + T(x_j), F(x_j)\Big),$$

which is linear in the semiring with operations $\oplus = \max$ and $\odot = +$, where $F(x_j) = t_0 + T(x_j)$ for minimal x_j (i.e., for x_j not preceded by any operations) and $F(x_j) = t_0$ otherwise. Thus, the working time of the parallel algorithm can be found by solving the shortest path problem on a cycle-free graph.

A more interesting problem arises for the case in which the processor set $P = \{p_1, \dots, p_m\}$ is small, that is, $m \ll n$. Then the computation scheduling problem is to construct a timetable S, i.e., a pair of mappings ($f: X \to \mathbb{R}$, $y: X \to P$), where f specifies the starting time for each operation and g specifies the processor-operation assignment. The timetable should be consistent with the given order on X and should give an optimum of some performance criterion, e.g., of the time by which all operations are finished.

Another type of problems deals with the construction of the most efficient parallel algorithms for a concrete problem (or a class of similar problems) on a processor with fixed structure, e.g., a matrix processor. The analysis of these problems is deeply connected with the underlying algebraic structure of an idempotent semiring or with more general structures containing three operations: $\max, \min$, and $+$. For the discussion of these problems, we refer the reader to [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**].

Mathematically similar problems appear in production scheduling. Suppose that in a workshop having a number of machines $M = \{M_1, \dots, M_m\}$ one should carry out a number of jobs $J = \{J_1, \dots, J_j\}$, each of which must be processed on various M_i before yielding a finished product. The times t_{ik} of processing the job J_i on the machine M_k are given. A schedule is a matrix of prescribed starting times T_{ik} of carrying out the job J_i on the machine M_k such that no time intervals associated to the same machine or to the same job ovelap. The scheduling problem consists in determining a schedule that optimizes a given criterion, for instance, the maximum completion time, often called the makespan. Usually, the following three main classes of workshops are considered: 1) flow-shop, where each product follows the same sequence of machines (although some mashines may be skipped); 2) job-shop, where for each job the sequence in which the machines are visited is fixed, but is not necessarily the same for various jobs; 3) open-shop, where each job must still be processed on a fixed set of machines but the order is not fixed. Let us consider the workshop of the flow-shop type: each J_i should be carried out successively on the machines $M_1, \dots, M_m$ (note that $t_{ik} = 0$ means that the machine M_k is to be skipped when carrying out the job J_i), and we suppose that there are no resetting times for machines when they switch from one job to another. Let us also assume that the order in which the jobs are processed on each machine is fixed; namely, on each M_k the jobs $J_1, \dots, J_j$ are successively carried out. To find an optimal schedule in the sense of minimal makespan means to find a matrix $T = T_{ik}$ of earliest times when the machine M_k can start the job J_i. To have a well-posed problem, we also need to have a given input $(j+m)$-vector $\mathrm{In} = (\mathrm{In}^J, \mathrm{In}^M)$, where $\mathrm{In}^J = (\mathrm{In}^J_1, \dots, \mathrm{In}^J_j)$ and $\mathrm{In}^M = (\mathrm{In}^M_1, \dots, \mathrm{In}^M_m)$ are the times when the jobs and the machines, respectively, are available. The result of the work is described by the output $(j+m)$-vector $\mathrm{Out} = (\mathrm{Out}^j, \mathrm{Out}^M)$, where $\mathrm{Out}^J = (\mathrm{Out}^J_1, \dots, \mathrm{Out}^J_j)$ and $\mathrm{Out}^M = (\mathrm{Out}^M_1, \dots, \mathrm{Out}^M_m)$ denote the times by which, respectively, the jobs have been carried out and the machines have finished their work. Obviously, the times T_{ik} satisfy the system of evolution equations

$$T_{ik} = \max(t_{i,k-1} + T_{i,k-1}, t_{i-1,k} + T_{i-1,k})$$

with the initial conditions

$$T_{i1} = \mathrm{In}_i^J, \quad T_{1k} = \mathrm{In}_k^M,$$

and the output vector is given by formulas

$$\mathrm{Out}_i^J = t_{im} + T_{im}, \quad \mathrm{Out}_k^M = t_{jk} + T_{jk}.$$

This system is linear in the algebra with operations $\oplus = \max$, $\odot = +$, and we can write $\mathrm{Out} = P\,\mathrm{In}$, where the transition matrix P can be calculated by the methods described above. Note also that this system coincides with that arising in the analysis of the operations of the Kung matrix processor [**?**, **?**]. Let us now suppose that our workshop must process a series of identical sets of jobs $J(k)$, $k = 1, 2, \dots$ If each machine after finishing the kth series can immediately start working with the $(k+1)$st series, then $\mathrm{In}(k) = \mathrm{Out}(k-1)$. Therefore, $\mathrm{In}(k) = P^k\mathrm{In}(1)$, and the solution of the scheduling problem considered determines the performance of the system. In the model considered, the processing times t_{ik} were constant. A model with variable processing times is considered from the point of view of idempotent algebra in [**?**]; see also [**?**] for the analysis of some examples of job-shop type scheduling.

8. Queueing system with finite capacity. The idea of this example is taken from [**?**]. In contrast to the preceding example, the evolution of the system is described by an implicit linear equation. Let us consider n servers S_i, $i = 1, \dots, n$. Each customer is to be successively served by all servers $S_1, \dots, S_n$. The times $t_j(k)$ necessary for S_j to serve the kth customer, $k = 1, 2, \dots$, are given. The input data are given by the instants $u(k)$ at which the kth customer arrives into the buffer associated with S_1. There are no buffers between the servers. As a consequence, if a server S_i, $i = 1, \dots, n-1$, has finished serving the kth customer but S_{i+1} is still busy with the $(k-1)$st customer, then S_i cannot start serving a new customer but has to wait until S_{j+1} is free. It is also assumed that the traveling times between the servers are zero. Let $x_i(k)$ denote the instant at which S_i begins to serve the kth customer. Then, obviously,

$$x_i(k+1) = \max(t_{i-1}(k+1) + x_{i-1}(k+1), t_i(k) + x_i(k), t_{i+1}(k-1) + x_{i+1}(k-1))$$

for $i = 1, \dots, n-1$, where it is assumed that $t_{-1}(k) + x_{-1}(k) = u(k)$ and $t_{n+1}(k) + x_{n+1}(k) = -\infty$. We can rewrite this equation in the matrix form

$$x(k+1) = A(k+1)x(k+1) \oplus B(k)x(k) \oplus C(k-1)x(k-1) \oplus Du(k+1),$$

where $\oplus = \max$,

$$A(l)_{i+1,i} = B(l)_{ii} = C(l)_{i-1,i} = t_i(l), \quad D_{11} = 0,$$

and the other entries of the matrices $A(l)$, $B(l)$, $C(l)$, and D are equal to $-\infty$ for all l. Thus, we have obtained an implicit equation for $x(k+1)$. To reduce this equation to an explicit form, one must first find the general solution of the stationary equation $X = A(k+1)X \oplus F$ with general F.

9. Petri nets and timed event graphs. The models presented here can be viewed as a far-going generalization of the previous example. For complete discussion of this topic, we refer the reader to the excellent recent book [**?**]; here we give only the main definitions and ideas. Let us recall first that a graph $\Gamma = (V, E)$ with node set V and arc set E is called a *bipartite graph* if the set V is the union of two disjoint subsets P and T such that there are no arcs between any two nodes of P as well as between any two nodes of T. In the literature on Petri nets, the elements of P are called *places* and the elements of T are called *transitions*. In the usual models, places represent conditions and transitions represent events. If $p \in P$, $t \in T$, and $(p, t) \in E$ (or $(t, p) \in E$), then p is called an *upstream* (respectively, *downstream*) place for the transition t. A function $\mu : P \mapsto \mathbb{Z}_+$ (or, equivalently, a $|P|$-vector $\mu = (\mu_1, \dots, \mu_{|P|})$ of nonnegative integers) is called a *marking* of the bipartite graph $\Gamma = (P \cup T, E)$. One says that the place p_i is marked with μ_i *tokens*. By definition, a *Petri net* is a bipartite graph equipped with some marking. In the standard graphic representation of Petri nets, places are drawn as circles and transitions as bars (or rectangles). Moreover, the number of dots placed in each circle is equal to the number of tokens marking the corresponding place. The dynamics of a Petri net is defined as follows. A transition t is said to be *enabled* if each of its upstream places contains at least one token, and the *firing* of an enabled transition t is, by definition, the marking transformation $\mu \mapsto T_t\mu$ described by the formula

$$(T_t\mu)_j = \begin{cases} \mu_j - 1 & \text{if } (p_j, t) \in E, \quad (t, p_j) \notin E \\ \mu_j + 1 & \text{if } (p_j, t) \notin E, \quad (t, p_j) \in E \\ \mu_j & \text{otherwise.} \end{cases}$$

A Petri net is said to be *timed* if associated with each transition is some firing time, i.e., the duration of firing (when it occurs), and associated with each place is some holding time, i.e., the time a token must spend in the place before contributing to the downstream transitions. These times need not be constant in general (they can depend on the number of firings of a transition or on some other parameters). A timed Petri net is called a *timed event graph* (TEG) if each place has exactly one upstream and one downstream transition. TEGs proved to be a very convenient tool in modeling a wide class of discrete event systems (DES). Assumung that each transition starts firing when enabled, we can now define the state variable $x_j(k)$, $j = 1, ..., |T|$, of a TEG (or, generally, of a timed Petri net) as the instant at which the transition t_j starts firing for the kth time. It was proved in [?] that for TEGs with some additional reasonable properties there exists an $M \in \mathbb{N}$ and matrices $A(k, k-j)$, $j = 0, \ldots, M$, $k \in \mathbb{N}$, such that

$$x(k) = \oplus_{j=0}^{M} A(k, k-j) x(k-j), \quad k = M, M+1, ...,$$

where the matrix multiplication is understood in the sense of the algebra with operations $\oplus = \max$ and $\odot = +$.

In closing, let us indicate some other recent developments in idempotent algebra (this is by no means a complete survey). Many of them are presented in [?]. The systems with three basic operations $\max, \min, +$ have been investigated in the series of works by G. J. Olsder and J. Gunawardena, see [?, ?] and references therein. Some results in this direction obtained from the point of view of game theory are given in [?]; see also Chapter 2. D.Cofer and V. Garg [?, ?] applied idempotent algebra to the solution of control problems for timed event graphs. A program of hardware and software design in computer engineering based on the correspondence principle is given in the recent paper [?] by G. L. Litvinov and V. P. Maslov, where a short survey of the development of idempotent analysis can be found (also, see [?]). Furthermore, there is an interesting field of applications of the theory of the idempotent semiring of integers, the so-called tropical semiring (e.g., see [?, ?, ?, ?] and references therein). There is a series of papers dealing with the applications of idempotent algebra to the investigation of stochastic discrete event systems modeled by stochastic Petri nets or, more precisely, by stochastic event graphs (SEG). The discussion of the main results in this direction can be found in the work of F. Baccelli [?] and J. Mairesse [?], where, in particular, some analogs of ergodic theorems for products of random matrices are obtained. Other stochastic applications of the (appropriately generalized) idempotent algebra, namely, applications to Markov processes and stochastic games, can be found in [?]; see also Chapter 2. Returning to the deterministic case, let us note that some new convergence properties of the iterates of linear operators in the $(\max, +)$-algebra were obtained by R. D. Nussbaum [?], whose work was motivated by problems of statistical mechanics, where the investigation of such iterates also proved to be of importance [?]. Finally, let us note the work of S. Gaubert [?, ?, ?] on rational series in idempotent algebras with applications to the dynamics of TEGs and a series of works by E. Wagneur on the classification of finite-dimensional semimodules [?, ?, ?], as well as his new paper in [?]. In the latter book, many other new interesting developments, as well as a historical survey, can be found.

1.3. The Main Theorem of Idempotent Analysis

In this section, on the basis of the results from [**?**, **?**, **?**, **?**], we prove a theorem describing the structure of endomorphisms of the space of continuous functions ranging in an idempotent semigroup. Recently, M. Akian [**?**] has generalized some of these results to the case of semirings whose topology is not defined by a metric. First, let us recall some definitions of the theory of ordered sets.

A set M is said to be (nonstrictly) *partially ordered* if it is equipped with a reflexive, transitive, antisymmetric relation $\leq$ (recall that antisymmetry implies that $a \leq b \ \wedge \ b \leq a \implies a = b$). We say that M is directed *upward* (respectively, *downward*) if $\forall\, a, b \in M \ \exists\, c : a \leq c \ \wedge \ b \leq c$ (respectively, $a \geq c \ \wedge \ b \geq c$). An element $c \in M$ is called an *upper bound* of a subset $\Pi \subset M$ if $c \geq p \ \forall p \in \Pi$; c is the *least upper bound* of Π if c is an upper bound of Π and $c \leq c'$ for any other upper bound c'. The lower bounds and the greatest lower bound are defined similarly. A subset of M is said to be *bounded above* (respectively, *below*) if it has an upper (respectively, a lower) bound. A partially ordered set M is called a *complete lattice* if each subset bounded above (respectively, below) has the least upper bound (respectively, the greatest lower bound). These bounds will be denoted by the usual symbols sup and inf.

A *net* (*Moore–Smith sequence*) $\{x_\alpha\}_{\alpha\in I}$ in an arbitrary set S is a mapping of an indexing set I directed upward into S. Usual sequences are special case of nets for $I = \mathbb{N}$. Let S be a topological space. A point $x \in S$ is called the *limit* (respectively, a *limit point*) of a net $\{x_\alpha\}_{\alpha\in I}$ in S if for each neighborhood $U \ni x$ there exists an $\alpha \in I$ such that $x_\beta \in U$ for all $\beta \geq \alpha$ (respectively, for each $\alpha \in I$ there exists a $\beta \geq \alpha$ such that $x_\beta \in U$). The convergence of nets in a set S uniquely determines the topology on this set, and moreover, the following conditions are satisfied:

1) S is a Hausdorff space $\iff$ each net in S has at most one limit;

2) a subset $\Pi \subset S$ is closed $\iff$ Π contains all limit points of each net all of whose elements lie in Π;

3) a function $f\colon T \to S$, where T and S are topological spaces, is continuous $\iff$ for each convergent net $\{x_\alpha\}_{\alpha\in I}$ in T, the net $\{f(x_\alpha)\}_{\alpha\in I}$ is convergent in S to $f(\lim\{x_\alpha\}_{\alpha\in I})$.

A statement concerning a net $\{x_\alpha\}_{\alpha\in I}$ is said to be *eventually true* if there exists an $\alpha' \in I$ such that this statement is true for the net $\{x_\alpha\}_{\alpha\geq\alpha'}$.

A net $\{x_\alpha\}_{\alpha\in I}$ is called a subnet of a net $\{y_\beta\}_{\beta\in J}$ if there exists a function $F\colon I \to J$ with the following properties: $x_\alpha = y_{F(\alpha)}$ for each $\alpha \in I$, and for each $\beta' \in J$ there exists an $\alpha' \in I$ such that $\alpha > \alpha'$ implies $F(\alpha) > \beta'$ (in other words, $F(\alpha)$ is eventually greater than β' for any $\beta' \in J$).

A space K is compact if and only if each net in K contains a convergent subnet.

In spaces with the first countability axiom (for example, in metric spaces) the topology is uniquely determined by the convergence of sequences.

If a partially ordered set M is a complete lattice, then there is a natural convergence on M. If a net $\{x_\alpha\}_{\alpha\in I}$ in M is eventually bounded, then we define its upper and lower limits by the formulas

$$\overline{\lim}\, x_\alpha = \inf_\alpha \sup_{\beta\ge\alpha} x_\beta, \qquad \underline{\lim}\, x_\alpha = \sup_\alpha \inf_{\beta\ge\alpha} x_\beta.$$

Obviously, $\overline{\lim}\, x_\alpha \ge \underline{\lim}\, x_\alpha$. If this inequality is actually an equality, then the net $\{x_\alpha\}$ is said to be convergent and the common value of both sides of this equality is called the limit of $\{x_\alpha\}$. Thus, we have defined a topology on M with the following properties:

1) All segments $[a,b] = \{x : a \le x \le b\}$ are closed sets.

2) Any nondecreasing (respectively, nonincreasing) net $\{x_\alpha\}_{\alpha\in I}$ bounded above (respectively, below) has the limit $\lim x_\alpha = \sup x_\alpha$ (respectively, $\inf x_\alpha$).

The topology thus introduced specifies a strict order $<$ on M by the formula

$$a < b \iff b \in \operatorname{int}\{c : c \ge a\},$$

where $\operatorname{int} C$ is the interior of a set C.

Recall that a strict partial order on a set is an antireflexive transitive binary relation on this set.

As was noted in §1.1, there is a one-to-one correspondence between associative commutative idempotent operations on an arbitrary set M and partial orders on M such that each two-point set has the greatest lower bound. This correspondence is given by the formula $a \oplus b = \inf(a,b)$. To obtain nontrivial results about function semimodules on infinite sets, it is convenient to modify the definition of an idempotent metric semigroup given in §1.1. In what follows by an idempotent metric semigroup we mean an idempotent semigroup M endowed with a metric ρ such that M is a complete lattice with respect to the order $\le$ corresponding to the operation $\oplus$ and the metric ρ is consistent with the order in the following sense:

M1. The order and the metric define the same topology; that is, for any net $\{x_\alpha\}_{\alpha\in I}$ in M the existence of $\lim x_\alpha = c$ in the topology induced by the partial order is equivalent to the relation $\rho(x_\alpha, c) \to 0$.

M2. Local boundedness: all balls $B_R(a) = \{y \in M : \rho(y,a) \le R\}$ are order-bounded below.

M3. Monotonicity of the metric:

$$a \le b \implies \rho(a,b) = \sup_{c\in[a,b]} \sup(\rho(a,c), \rho(c,b)).$$

M4. The minimax property of the metric:

$$\rho(a\oplus b, c\oplus d) \le \max(\rho(a,c), \rho(b,d)).$$

We already know that the last condition implies the uniform continuity of the $\oplus$-addition in the metric space M. For all examples of semigroups in §1.1, conditions M1–M4 are satisfied.

The following criterion holds for the first two conditions.

Lemma 1.1 *If all balls $B_R(x)$ are compact in the topology defined by the metric, then*

a) M4 *implies* M2;

b) *under conditions* M4 *and* M3*, any net convergent with respect to the order is convergent with respect to the metric*;

c) *suppose that condition* M4*, as well as the following minimax condition, is satisfied:*

M5. $\rho(\sup(a,b),\sup(c,d)) \le \max(\rho(a,c),\rho(b,d))$ *with respect to the operation* sup.

Then convergence with respect to the metric implies convergence with respect to the order.

Proof. a) Suppose that M4 is satisfied and the ball $B_R(a)$ is compact. Consider the set I of finite tuples $\alpha = \{a_1,\dots,a_n\}$ of points in M. The set I is directed by inclusion. Consider the net

$$x_\alpha = a_1 \oplus \dots \oplus a_n = \inf(a_1,\dots,a_n)$$

indexed by I. It follows from condition M4 that $x_\alpha \in B_R(a)$ for all α, and consequently, by the above compactness criterion, there exists a subnet $\{y_\beta\}_{\beta\in J}$ of the net $\{x_\alpha\}_{\alpha\in I}$ such that $\{y_\beta\}$ is convergent with respect to the metric to some point $b \in B_R(a)$. It follows from the definition of a subnet and from the continuity of the $\oplus$ operation that $b \le c$ for each $c \in B_R(a)$, and consequently, $B_R(a)$ is bounded.

b) Let conditions M3 and M4 be satisfied. First, let us prove that any net $\{x_\alpha\}_{\alpha\in I}$ that is nonincreasing and bounded below is convergent with respect to the metric to the greatest lower bound $\inf\{x_\alpha\}$. Indeed, since $\{x_\alpha\}$ is bounded, it follows that all elements $\{x_\alpha\}$ are contained in some ball, which is compact by condition. Consequently, from $\{x_\alpha\}_{\alpha\in I}$ we can extract a convergent subnet $\{y_\beta\}_{\beta\in J}$. Just as in a), we obtain $c = \lim y_\beta \le x_\alpha$ for all $\alpha \in I$. Hence, $c \le \inf\{x_\alpha\}$. Since $\rho(y_\beta, c) \to 0$, it follows that $c = \inf\{x_\alpha\}$. Since the metric is monotone, it follows that $\rho(x_\alpha, c) \to 0$, that is, $c = \inf\{x_\alpha\} = \lim x_\alpha$.

Similarly, it can be proved that any nonincreasing net $\{x_\alpha\}_{\alpha\in I}$ is convergent with respect to the metric to the least upper bound $\sup\{x_\alpha\}$. Now let a bounded net $\{x_\alpha\}$ satisfy the convergence condition in the sense of the partial order:

$$\inf_\alpha \sup_{\beta\ge\alpha} x_\alpha = \sup_\alpha \inf_{\beta\ge\alpha} x_\beta = c.$$

Then, as was shown in the preceding,

$$\rho(\sup_{\beta\ge\alpha}\{x_\beta\}, c) \to 0 \quad \text{and} \quad \rho(\inf_{\beta\ge\alpha}\{x_\beta\}, c) \to 0.$$

By virtue of property M3, monotonicity of the metric, and the triangle inequality, we have

$$\rho(x_\alpha, c) \le \rho\Big(x_\alpha, \sup_{\beta\ge\alpha} x_\beta\Big) + \rho\Big(\sup_{\beta\ge\alpha} x_\beta, c\Big)$$
$$\le \rho\Big(\inf_{\beta\ge\alpha} x_\beta, \sup_{\beta\ge\alpha} x_\beta\Big) + \rho\Big(\sup_{\beta\ge\alpha} x_\beta, c\Big) \to 0,$$

that is, $\lim x_\alpha = c$.

c) Let conditions M4 and M5 be satisfied, and let $\{x_\alpha\}_{\alpha\in I}$ be a metric-convergent net in M such that $\rho(x_\alpha, c) \to 0$ for some $c \in M$. Then the net $\{x_\alpha\}$ is eventually bounded and hence lies in a compact set. As in the proof of a), from M4 and M5 we conclude that if $x_\alpha \in B_R(x)$ for $\alpha \ge \alpha'$, then for $\alpha \ge \alpha'$ the elements of the decreasing net $\sup_{\beta\ge\alpha} x_\beta$ and the increasing net $\inf_{\beta\ge\alpha} x_\beta$ also belong to this ball. It follows that these nets are also convergent to $c = \lim x_\alpha$ with respect to the metric. It follows from this fact and from the continuity of the operation $\oplus$ that

$$\inf_\alpha \sup_{\beta\ge\alpha} x_\beta \le c \le \sup_\alpha \inf_{\beta\ge\alpha} x_\beta, \qquad \alpha > \alpha',$$

that is, $\overline{\lim}\, x_\alpha \le c \le \lim x_\alpha$. Since $\underline{\lim}\, x_\alpha \le \overline{\lim}\, x_\alpha$ for any net, we conclude that $\overline{\lim}\, x_\alpha = \underline{\lim}\, x_\alpha = c$.

We shall also need the following properties of order on an idempotent metric semigroup.

Lemma 1.2 a) *For each ball $B_R(a)$, we have* $\inf\{y \in B_R(a)\} \in B_R(a)$, *that is, each ball contains its greatest lower bound*;

b) *if a is not a local minimum, that is, if $a \ne \inf\{y \in U\}$ for any neighborhood U of a, then in each neighborhood of a there exists a point c such that $c < a$*;

c) $a < b \le c \implies a < c$.

Proof. a) For each ball $B_R(a)$, consider the net $\{x_\alpha\}$ constructed in the proof of item a) in Lemma 1. By the minimax property of the metric, all elements x_α lie in $B_R(a)$. The net $\{x_\alpha\}$ is nonincreasing and hence converges to its greatest lower bound, which lies in the ball since the latter is closed.

b) This readily follows from a).

c) Let $B_R(b)$ be a ball such that $a \le y$ for each $y \in B_R(b)$. By the minimax property of the metric, $\inf\{b, z\} \in B_R(b)$ for any $z \in B_R(c)$. Thus, by analogy with a), we see that $a \le \inf\{b, B_R(c)\}$.

The proof is complete.

An idempotent metric semigroup is said to be *connected* (more precisely, *order-connected*) if all segments $[a, b] \in M$ are connected sets.

Lemma 1.3 *Let M be a connected idempotent metric semigroup. Then each segment $[a,b] \subset M$ is arcwise connected (that is, there exists a continuous mapping $f\colon [0,1] \to [a,b]$, where $[0,1] \subset \mathbb{R}$, such that $f(0)=a$ and $f(1)=b$).*

Proof. Let $a \le b$ and $a \ne b$. Then the set

$$\Pi_{[a,b]} = \{x \in [a,b] : \rho(a,x) = \rho(x,b)\}$$

of midpoints of the segment $[a,b]$ is nonempty. Indeed, otherwise the sets

$$A = \{x \in [a,b] : \rho(a,x) < \rho(x,b)\},$$
$$B = \{x \in [a,b] : \rho(a,x) > \rho(x,b)\}$$

would be nonempty open-closed subsets in $[a,b]$ such that $A \cap B = \emptyset$ and $A \cup B = [a,b]$, which is impossible by definition, since $[a,b]$ is connected.

Let us now construct a mapping $f\colon D \to [a,b]$ of the set $D \subset [0,1] \subset \mathbb{R}$ of dyadic numbers as follows. Set $f(0)=a$ and $f(1)=b$. Next, we set $f(1/2)$ to be an arbitrary element of $\Pi_{[a,b]}$. Furthermore, $f(1/4)$ and $f(3/4)$ are defined as arbitrary elements of the sets $\Pi_{[a,f(1/2)]}$ and $\Pi_{[f(1/2),b]}$. Thus, we specify the value of f by induction at any point of the form $k/2^n \in D$, $k,n \in \mathbb{N}$, $k < 2^n$. Obviously, f is continuous on D and extends to be a continuous function on the closure $\overline{D} = [0,1]$; the continuation is given by the formula

$$f(x) = \sup\{f(d) : d \in D,\, d \le x\} = \inf\{f(d) : d \in D,\, d \ge x\},$$

where $x \in [0,1]$ is arbitrary. The proof is complete.

Let us consider two examples showing that the minimax property M4 is important and nontrivial: it can be violated for apparently quite natural metrics on ordered spaces, and if it is violated, then most statements in Lemmas 1.1 and 1.2 fail to be true. First, note that this property is not valid for the usual Euclidean metric on $\mathbb{R}^n_+$ equipped with the Pareto partial order, but is valid for the metric $\rho(x,y) = \max_{i=1,\dots,n} |x^i - y^i|$, $x = \{x^i\}$, $y = \{y^i\}$, which is a natural metric on this idempotent semigroup. Furthermore, the minimax property is not inherited by subsets of ordered sets. In the Pareto partially ordered unit square $Q \subset \mathbb{R}^2_+$, consider the union $C \subset Q$ of the diagonals equipped with the inherited metric, inherited order, and the corresponding operation $\oplus$ (see Fig. 1 (a)). The sequence $z_k = (x_k, y_k)$, where $x_k = 1 - y_k$ and $y_k = 1/2^k$ converges in the metric to $(1,0)$, but its upper and lower limits (in the sense of order) are equal to $(1,1)$ and $(0,0)$, respectively. It is easy to see that property M4 is violated in this example, and so is property M1, even though all balls are compact sets.

On the other hand, the violation of property M4 for the subset hatched in Figure 1 (b) implies that property c) in Lemma 1.2 is not valid for the points a, b, and c.

Let us now proceed to the study of functions from a topological space X to an idempotent metric semigroup M. For further reference, we state two obvious properties in the following lemma.

Fig. 1. (a) (b)

Lemma 1.4 a) *$\rho\Big(\inf_x h(x), \inf_x g(x)\Big) \le \sup_x \rho(h(x), g(x))$ for any bounded functions h, $g\colon X \to M$;*

b) *if a net $\{f_\alpha\colon X \to M\}_{\alpha \in I}$ of continuous functions is monotone increasing (or decreasing) and pointwise convergent to a continuous function f, then it converges to f uniformly on each compact set.*

The partial order structure permits us to give two definitions of semicontinuity for functions $f\colon X \to M$; both definitions generalize the corresponding notion for real-valued functions.

Definition SC1 A function $f\colon X \to M$ is said to be *lower* (respectively, *upper*) *semicontinuous* at a point $x \in X$ if $\underline{\lim}\, f(x_\alpha) \ge f(x)$ (respectively, $\overline{\lim}\, f(x_\alpha) \le f(x)$) for any net $\{x_\alpha\}_{\alpha \in I}$, $x_\alpha \in X$, convergent to x.

Definition SC2 A function $f\colon X \to M$ is said to be *lower* (respectively, *upper*) *semicontinuous* at a point $x \in X$ if the inequality $f(x) > a$ (respectively, $f(x) < a$) implies that $f(y) > a$ (respectively, $f(y) < a$) for y in some neighborhood U of x in X.

Lemma 1.5 *If M is an idempotent metric semigroup, then definitions* SC1 *and* SC2 *of lower semicontinuity are equivalent.*

Proof. 1) Assume that SC2 is violated. Then there exists an $a < f(x)$ and a net $\{x_\alpha\}_{\alpha \in I}$ convergent to x such that the inequality $f(x_\alpha) > a$ is violated for each $\alpha \in I$. Then the inequality $\underline{\lim}\, f(x_\alpha) > a$ is obviously violated as well, and the same is true of the inequality $\underline{\lim}\, f(x_\alpha) \ge f(x)$ by virtue of item c) in Lemma 1.2.

2) Suppose that the requirement of Definition SC2 is satisfied but the requirement of Definition SC1 is not satisfied for some $x \in X$. Let us take a net $\{x_\alpha\}_{\alpha \in I}$ convergent to x and such that the inequality $\underline{\lim}\, f(x_\alpha) \ge f(x)$ is not valid. By using item b) in Lemma 1.2, we can find a $c < f(x)$ such that the inequality $\underline{\lim}\, f(x_\alpha) \ge c$ also fails. But this contradicts the fact that, by Definition SC2, $f(x_\alpha) > c$ for α greater than some α'. The lemma is proved.

Throughout the remaining part of this section, M is assumed to be a connected idempotent metric semigroup.

Let $C_0^\infty(X,M)$ denote the space of functions $f\colon X\to M$ on a locally compact normal space X such that f tends to $\mathbb{0}$ at infinity; the convergence on $C_0^\infty(X,M)$ is defined by the uniform metric $\rho(f,g)=\sup_X \rho(f(x),g(x))$. Let $C_0(X,M)\subset C_0^\infty(X,M)$ be the subspace of continuous functions $f\colon X\to M$ with compact support $\operatorname{supp}_{\mathbb{0}} f=\overline{\{x : f(x)\neq \mathbb{0}\}}$; we say that a net is convergent in $C_0(X,M)$ if it is convergent in the uniform metric (i.e., in $C_0^\infty(X,M)$) and if for some compact set $K\subset X$ the supports of its elements eventually lie in K. The space $C_0(X,M)$, as well as $C_0^\infty(X,M)$, is an idempotent semigroup with respect to the pointwise idempotent addition, which will also be denoted by $\oplus$ and which is continuous on $C_0(X,M)$. Note, however, that the function space $C_0(X,M)$ is not a complete lattice in general. Let us introduce the semigroup L_{up} of functions $f\colon X\to M$ that can be represented as pointwise limits of monotone increasing nets in $C_0(X,M)$ with common compact support. It is easy to see that all functions in L_{up} are lower semicontinuous. Let us define characteristic functions of points by the formula

$$g_x^a(y)=\begin{cases} a & \text{if } y=x,\\ \mathbb{0} & \text{if } y\neq x\end{cases}$$

(here $x\in X$ and $a\in M$).

Lemma 1.6 *The characteristic functions g_x^a belong to L_{up}.*

Proof. Let K be the compact closure of some neighborhood of the point $x\in X$, and let I be the directed set of pairs (U,V) of neighborhoods of x such that $\overline{U}\subset V\subset K$ with the partial order $(U_1,V_1)\prec(U_2,V_2)\iff \overline{V}_2\subset U_1$. Since the segment $[a,\mathbb{0}]$ is connected in M, it follows from the Urysohn lemma that for each pair $(U,V)\in I$ there exists a function $h_{(U,V)}\in C_0(X,M)$ such that $h(y)\in[a,\mathbb{0}]$ for any y, $\operatorname{supp}_{\mathbb{0}} h\in V$, and $h|_U=a$. Obviously, the net $\{h_{(U,V)}\}$ is monotone increasing and converges to g_x^a. The lemma is proved.

We are now in a position to prove the main result of this chapter.

Theorem 1.3 *Let M be a connected idempotent metric semigroup, and let $m\colon C_0(X,M)\to M$ be a homomorphism. Then m can be extended by monotonicity and continuity to a mapping $L_{\mathrm{up}}\to M$, and so the mapping*

$$f\colon M\times X\to M,\qquad f(a,x)=m(g_x^a)$$

is well defined. Furthermore, f is jointly lower semicontinuous in (a,x) and is additive with respect to the first argument, that is,

$$f(a\oplus b,x)=f(a,x)\oplus f(b,x).$$

The homomorphism m can be reconstructed from f by the formula

$$m(h)=\inf_x f(h(x),x)\qquad \forall\, h\in C_0(X,M). \tag{1.20}$$

Proof. Let a net $\{h_\alpha\}_{\alpha\in I}$ of functions in $\mathbb{C}_{\mathbf{0}}(X,M)$ with common compact support be monotone increasing and converge to some $\varphi \in L_{\mathrm{up}}$. In this case we write $h_\alpha \nearrow \varphi$. Set

$$m(\varphi) = \lim m(h_\alpha) = \sup m(h_\alpha).$$

Let us show that this continuation of m to L_{up} is well defined. Indeed, suppose that $\{h_\alpha\}_{\alpha\in I} \nearrow \varphi$ and $\{h'_\beta\}_{\beta\in j} \nearrow \varphi$, where I and J are directed indexing sets. By Lemma 1.2, d) it follows from the continuity of m that

$$m(h_0) = \lim_{\beta\in J} m(h_{\alpha_0} \oplus h'_\beta) \le \lim_{\beta\in J} m(h'_\beta),$$

and consequently,

$$\lim_{\alpha\in I} m(h_\alpha) \le \lim_{\beta\in J} m(h'_\beta).$$

The reverse inequality can be proved in a similar way.

Let us prove that the mapping $f(a,x) = m(g_x^a)$ is lower semicontinuous. If $f(a,x) > c$, then it follows from Lemma 1.2, b) that there exists a $d < a$, a neighborhood U of x, and a function $h \in C_{\mathbf{0}}(X,M)$ such that $h|_U = d$, $h(y) \ge d$ for all y, and $m(h) > c$. Consequently, for any $y \in V$ and $b > d$ we have $f(b,y) \ge m(h) > c$, and $f(b,y) > c$ by Lemma 1.2, b).

The additivity of f with respect to the first argument is obvious. Let us prove formula (1.20). We denote the right-hand side in (1.20) by c. Obviously, $m(h) \le f(h(x),x)$ for any x, and consequently, $m(h) \le c$. By virtue of Lemma 1.2, c), to prove the reverse inequality, it suffices to establish the implication $d < c \implies m(h) \ge d$. By Lemma 1.2, b), for each $x \in \operatorname{supp}_{\mathbf{0}} h$ there exists an $h_x \in C_{\mathbf{0}}(X,M)$ such that $h_x(y) \ge h(y)$ for every y, $h_x(x) = h(x)$, and $m(h_x) > d$. Consider the set I of finite tuples $\{x_1,\dots,x_k\}$, $x_j \in X$, ordered by inclusion. The net $h_{x_1,\dots,x_k} = h_{x_1} \oplus \dots \oplus h_{x_k}$ indexed by I converges to h and is monotone decreasing; hence, the convergence is uniform. Since $m(h_{x_1},\dots,h_{x_k}) \ge d$, it follows that $m(h) \ge d$. The theorem is proved.

Remark 1.2 Another proof of formula (1.20), which does not appeal to nets, can be obtained on the basis of a generalization of Tietze's extension theorem to M-valued functions. Namely, according to that theorem, for any point x there exists a neighborhood $U(x)$ and a function $h_x \in C_{\mathbf{0}}(X,M)$ such that $h_x \ge h$, $h_x|_U \equiv h|_U$, and $m(h_x) > d$. We can find a finite cover of the compact set $\operatorname{supp}_{\mathbf{0}} h$ by neighborhoods $U(x_k)$, $k = 1,\dots,n$. Then $h = h_{x_1} \oplus \dots \oplus h_{x_n}$, and consequently, $m(h) > d$.

Theorem 1.4 *Each continuous additive mapping* $B\colon C_{\mathbf{0}}(X,M) \to C_{\mathbf{0}}(X,M)$ *can be represented in the form*

$$(Bh)(y) = \inf_x b(x,y,h(x)), \tag{1.21}$$

where $b\colon X \times X \times M \to M$ *is some function such that* $b(x,y,\cdot\,)$ *is an endomorphism of the semigroup* M *for any given* $x,y \in X$.

Theorem 2 readily follows from Theorem 1.

Suppose that M is additionally equipped with a binary associative commutative operation $\odot$ with neutral element $\mathbb{1}$. Furthermore, suppose that $\odot$ distributes over $\oplus$, satisfies the condition $a \odot \mathbb{0} = \mathbb{0}$ $\forall a \in M$, and is uniformly continuous on bounded subsets of M. Then M will be called an *idempotent metric semiring*. In this case, the semigroup $C_{\mathbb{0}}(X, M)$ is actually a semimodule. The homomorphisms $C_{\mathbb{0}}(X, M) \to M$ will be called *linear functionals* on $C_{\mathbb{0}}(X, M)$. The set of linear functionals will be denoted by $C^*_{\mathbb{0}}(X, M)$ and called the *dual semimodule* of $C_{\mathbb{0}}(X, M)$.

Theorem 1.5 *Let M be an idempotent metric semiring. Then for any $m \in C^*_{\mathbb{0}}(X, M)$ there exists a unique lower semicontinuous function $f: X \to M$ such that*

$$m(h) = \inf_x f(x) \odot h(x) \qquad \forall h \in C_{\mathbb{0}}(X, M). \tag{1.22}$$

Conversely, any function $f: X \to M$ locally bounded below defines an element $m \in C^*_{\mathbb{0}}(X, M)$ by formula (1.22).

The proof of the first part of this theorem readily follows from Theorem 1.4, and the continuity of the functional (1.22) follows from the inequality

$$\rho\Big(\inf_x f(x) \odot h_1(x), \inf_x f(x) \odot h_2(x)\Big) \le \sup_x \rho(f(x) \odot h_1(x), f(x) \odot h_2(x)),$$

which, in turn, follows from Lemma 1.4, a) and from the uniform continuity of the operation $\odot$ on any bounded set.

The following corollary of Theorem 1.4 generalizes Proposition 1.1. Suppose that $A = \mathbb{R} \cup \{+\infty\}$ is the semiring described in Example 1.1, End is the semiring of endomorphisms of the semigroup A, and X is a compact set.

Theorem 1.6 *There is an isomorphism between the semiring of additive operators on $C(X, A)$ (i.e., mappings $B: C(X, A) \to C(X, A)$ such that $B(h \oplus g) = B(h) \oplus B(g)$ for any $h, g \in C(X, A)$) and the semiring of linear operators on $C(X, End)$ (i.e., endomorphisms of the semimodule $C(X, End)$). This isomorphism takes any operator B of the form* (1.21) *to the operator $\widetilde{B}$ given by the formula*

$$(\widetilde{B}h)(x, m) = \inf_y b(x, y, h(y, m)), \qquad h \in C(X, \mathrm{End}).$$

Let us now describe endomorphisms of convolution semirings. Let CS be the convolution semiring from Example 1.8, where $X \subset \mathbb{R}^n$ and $A = \mathbb{R} \cup \{\infty\}$.

Theorem 1.7 *Each endomorphism of CS has the form*

$$(Bh)(x) = c \cdot \inf\{h(y) : y \in \xi^{-1}(x)\},$$

where $c > 0$ is a positive constant and $\xi: \mathbb{R}^n \to \mathbb{R}^n$ is an operator linear in the usual sense.

The proof of this theorem, as well as of its modification to the number semiring $(\mathbb{R} \cup \{\infty\}, \oplus = \min, \odot = \max)$, is given in [**?**]. The representation (1.21) of endomorphisms is the key point in this proof.

Let us generalize the main result of this section to the infinite-dimensional case, in which X is not a locally compact space. The corresponding linear operators naturally arise as Bellman operators in infinite-dimensional optimization problems. For example, in the theory of controlled quantum-mechanical systems, X is a Hilbert space. To extend the preceding result to this case, it is necessary to use the pointwise convergence topology in function semimodules.

Let X be a totally regular topological space, and let $C_p(X, M)$ be the space (semimodule) of continuous functions $X \to M$ equipped with the pointwise convergence topology. Here M is an idempotent metric semiring.

Theorem 1.8 ([?, ?]) *Let $m: C_p(X, M) \to M$ be a continuous homomorphism of semimodules. Then there exists a function $f: X \to M$ such that Eq.* (1.22) *is valid.*

The proof is close to that of Theorem 1.5 and is even simpler, since we need not keep track of uniform convergence. A similar result is valid for the space of bounded A-valued continuous functions.

1.4. Idempotent Measures, Integrals, and Distributions. Duhamel's Principle

For simplicity, all subsequent constructions are carried out for the number semiring $A = \mathbb{R} \cup \{+\infty\}$ with semigroup operations $\oplus = \min$ and $\odot = +$ and neutral elements $\mathbb{0} = +\infty$ and $\mathbb{1} = 0$, respectively, equipped with the metric $\rho_{\exp}(a, b) = |\exp(-a) - \exp(-b)|$ and with the natural ordering. The modifications necessary in the more general situation are usually quite obvious. Let us recall our notation. Let X be a separable locally compact topological space. By $C_0^\infty(X, A)$ we denote the space of continuous functions $f: X \to A$ that tend to $\mathbb{0} = +\infty$ at infinity (that is, for any $\varepsilon > 0$ there exists a compact set $K \subset X$ such that $\rho(f(x), \mathbb{0}) < \varepsilon$ for all $x \notin K$); furthermore, $C_0(X, A) \subset C_0^\infty(X, A)$ is the subspace of functions f with compact support

$$\operatorname{supp}_{\mathbb{0}} f = \overline{\{x : f(x) \notin \mathbb{0}\}}.$$

The topology on $C_0^\infty(X, A)$ and $C_0(X, A)$ was introduced in §1.3. These function spaces are topological semimodules with respect to pointwise addition $\oplus$ and pointwise multiplication $\odot$ by elements of A. The dual modules $C_0^\infty(X, A)^*$ and $C_0(X, A)^*$ are defined as semimodules of continuous homomorphisms (A-linear functionals) $C_0^\infty(X, A) \to A$ and $C_0(X, A) \to A$, respectively. The semimodule $C_0^\infty(X, A)^*$ possesses a natural metric; namely, if $m_1, m_2 \in C_0^\infty(X, A)^*$, then

$$\rho(m_1, m_2) = \sup\{\rho(m_1(h), m_2(h)) : h \in C_0^\infty(X, A), \rho(h, \mathbb{0}) \le 1\}.$$

If X is a compact set, then both $C_{\mathbb{0}}^{\infty}(X, A)$ and $C_{\mathbb{0}}(X, A)$ are the same semimodule $C(X, A)$ of all continuous A-valued functions on X.

In what follows, we usually omit the letter A in the notation of function semimodules, since the semiring A is fixed.

Item (a) of the following theorem is a special case of Theorem 1.5; the other items are direct corollaries of Theorem 1.5.

Theorem 1.9 *Let us assign a functional* $m_f: C_{\mathbb{0}}(X) \to A$ *to any locally bounded* (*in the metric* ρ) *function* $f: X \to A$ *by setting*

$$m_f(h) = \inf_x f(x) \odot h(x). \tag{1.23}$$

Then

(a) *The mapping* $f \mapsto m_f$ *is an isomorphism of the semimodule of lower semicontinuous functions onto the semimodule* $C_{\mathbb{0}}^{*}(X)$.

(b) *The functionals in* $(C_{\mathbb{0}}^{\infty}(X))^{*}$ *correspond to bounded functions* f.

(c) *The cited isomorphism is an isometry*; *that is,*

$$\sup_x \rho(f_1(x), f_2(x)) = \sup_h \rho(m_{f_1}(h), m_{f_2}h))$$

for any $m_{f_1}, m_{f_2} \in (C_{\mathbb{0}}^{\infty}(X))^{*}$, *where the supremum on the right-hand side is taken over the set* $\{h \in C_{\mathbb{0}}(X) : \rho(h, \mathbb{0}) \leq 1\}$.

(d) *The functionals* m_{f_1} *and* m_{f_2} *coincide if and only if the functions* f_1 *and* f_2 *have the same lower semicontinuous closures*; *that is,* $\operatorname{Cl} f_1 = \operatorname{Cl} f_2$, *where*

$$(\operatorname{Cl} f)(x) = \sup\{\psi(x) : \psi \leq f,\ \psi \in C(X, A)\}.$$

Remark 1.3 The assertion that any linear functional can be represented in the form (1.23) can be generalized to the semimodule of continuous functions having a limit at infinity. To this end, it suffices to use the one-point compactification of X. On the other hand, simple examples [**?**] show that Theorem 1.9 is not valid for the semimodule $C(X, A)$ of bounded continuous functions $f: X \to A$ even for $X = \mathbb{R}$.

The following idempotent analog of Fürstenberg's strict ergodicity theorem can readily be derived from Theorem 1.9.

Corollary 1.1 *The following properties of a homomorphism* $T: X \to X$ *of a metric compactum* X *are equivalent*:

(a) *the functional* $m_{\mathbb{1}} \in C(X)^{*}$ *is the only* T*-invariant functional*;

(b) $\displaystyle\lim_{n\to\infty} \bigoplus_{k=1}^{n} h(T^k x) = \min_{y \in X} h(y)$ *for any continuous function* $h: X \to A$ *and any* $x \in X$;

(c) *the homomorphism* T *is minimal; that is, the trajectory of each point is dense in* X.

The Riesz–Markov theorem in functional analysis establishes a one-to-one correspondence between continuous linear functionals on the space of continuous real functions with compact support on a locally compact space X and regular Borel measures on X. Let us discuss a similar correspondence in idempotent analysis.

An A-valued function μ defined on all subsets of X and completely additive (that is, satisfying the property

$$\mu\left(\bigcup_\alpha B_\alpha\right) = \bigoplus_\alpha \mu(B_\alpha) = \inf_\alpha \mu(B_\alpha)$$

for any family $\{B_\alpha\}$ of subsets of X) will be called an *A-measure* on X. Obviously, the set of A-measures on X is in a one-to-one correspondence with the set of functions $f\colon X \to A$ bounded below; for any function f the corresponding measure μ_f is given by

$$\mu_f(B) = \inf\{f(x) : x \in B\}.$$

Consequently, by Theorem 1.9, to each measure μ_f there corresponds a functional $m_f \in (C_{\mathbf{0}}^\infty(X))^*$. However, the inverse mapping is not single-valued, since various functions with the same lower semicontinuous closure define the same functional, but different measures. The set of all measures corresponding to a given functional m contains a minimal measure, which corresponds to the lower semicontinuous function f determining the functional m. Equation (1.23) specifies a continuation of m_f to the set of A-valued functions bounded on $\operatorname{supp}_{\mathbf{0}} f$. On analogy with conventional analysis, we say that such functions are integrable with respect to the measure μ_f and denote the values taken by m_f on these functions by the *idempotent integral*

$$m_f(h) = \int_X^\oplus h(x)\, d\mu_f(x) = \inf_x f(x) \odot h(x). \tag{1.24}$$

It turns out that many of the properties of the conventional integral hold in this situation as well.

1. The idempotent integral is linear, that is, for any $a_1, a_2 \in A$ and $h_1, h_2 \in C_{\mathbf{0}}^\infty(X)$ we have

$$\int_X^\oplus (a_1 \odot h_1(x) \oplus a_w \odot h_2(x))\, d\mu_f(x) = a_1 \odot \int_X^\oplus h_1(x)\, d\mu_f(x) \oplus a_2 \odot \int_X^\oplus h_2(x)\, d\mu_f(x).$$

2. The idempotent integral is the limit of Riemann sums. Let h be continuous on the unit cube $B \subset \mathbb{R}^n$, and let $B = \bigcup B_j$ be a partition of B into cubes of diameter $\leq \varepsilon$. Then

$$\int_B^\oplus h(x)\, d\mu_f = \lim_{\varepsilon \to 0} \bigoplus_j h(x_j) \odot \mu_f(B_j),$$

where x_j is an arbitrary point in B_j.

3. The idempotent integral is the limit of Lebesgue integral sums. For a function g bounded below on $\operatorname{supp}_{\mathbb{0}} f = \overline{\{x : f(x) \neq \mathbb{0}\}}$, we introduce a finite ε-net $\{\mathbb{0} = d_0 > d_1 > \cdots > d_n = d\}$ on the closed interval $[d, \mathbb{0}]$, where $d = \inf\{g(x) : x \in \operatorname{supp}_{\mathbb{0}} f\}$, and consider the sets $q_i = \{x : d_i \leq g(x) \leq d_{i-1}\}$. Then

$$\bigoplus_{i=1}^{n} d_i \odot \mu_f(q_i) \leq \int_X^{\oplus} g(x)\, d\mu_f(x) \leq \bigoplus_{i=1}^{n} d_{i-1} \odot \mu_f(q_i),$$

and both sides tend to the integral as $\varepsilon \to 0$.

4. The following idempotent analog of Fubini's theorem is valid: if $g \colon X_1 \times X_2 \to A$ is a bounded function and μ_{f_1} and μ_{f_2} are idempotent measures on X_1 and X_2, respectively, then the following iterated integrals coincide:

$$\int_{X_1}^{\oplus} \left(\int_{X_2}^{\oplus} g(x_1, x_2)\, d\mu_2(x_2) \right) d\mu_1(x_1)$$
$$= \int_{X_2}^{\oplus} \left(\int_{X_1}^{\oplus} g(x_1, x_2)\, d\mu_1(x_1) \right) d\mu_2(x_2).$$

All these assertions readily follow from the basic formula (1.24).

Let us now consider the general A-linear evolution equation in $C_{\mathbb{0}}(X)$. It has the form

$$\begin{aligned} \varphi_t(x) &= \varphi_0(x) \oplus \int_{(0,t]}^{\oplus} (L\varphi_\tau(x) \oplus \psi_\tau(x))\, d\mu_\tau, \\ \varphi_t(x)|_{t=0} &= \varphi_0(x), \end{aligned} \tag{1.25}$$

and is an idempotent analog of the integrated linear differential equation $\partial\varphi/\partial t = L\varphi$. In Eq. (1.25), μ_τ is an idempotent measure on the semiaxis, L is an endomorphism of the semimodule $C_{\mathbb{0}}(X)$ (or an A-linear operator on $C_{\mathbb{0}}(X)$), and $\psi_\tau(x)$ is a given family of functions.

Let us substitute the expression for φ given by this equation into its right-hand side and apply Fubini's theorem. By repeating this procedure n times, we obtain

$$\varphi_t = \bigoplus_{k=0}^{n} L^k \left(\varphi_0 \odot m_{0,t}^k \;\oplus \int_{(0,t]}^{\oplus} \psi_\tau \odot m_{\tau,t}^k\, d\mu_\tau \right)$$
$$\oplus L^{n+1} \int_{(0,t]}^{\oplus} \varphi_\tau \odot m_{\tau,t}^n\, d\mu_\tau,$$

where $m_{\tau,t}^0 = \mathbb{1}$ and

$$m_{\tau,t}^k = \underbrace{m_{\tau,t} \odot \cdots \odot m_{\tau,t}}_{k \text{ factors}}.$$

Passing to the limit as $n \to \infty$, we obtain the following assertion.

Theorem 1.10 *Equation* (1.25) *is solvable if and only if the series*

$$\varphi_t^* = \bigoplus_{k=0}^{\infty} L^k \left(\varphi_0 \odot m_{0,t}^k \oplus \int_{(0,t]}^{\oplus} \psi_\tau \odot m_{\tau,t}^k \, d\mu_\tau \right)$$

is convergent. In this case, φ_t^* *is a solution of Eq.* (1.25). *Moreover,* φ_t^* *is the neutral element in the set of all solutions; that is,* $\varphi_t \oplus \varphi_t^* = \varphi_t$ *for any solution* φ_t *of Eq.* (1.25).

On analogy with the finite-dimensional case (see §1.2), the solution φ_t^* of Eq. (1.25) will naturally be called the *Duhamel solution*, since it can be represented in the form

$$\varphi_t^* = G_t \varphi_0 \oplus \int_{(0,t]}^{\oplus} G_{\tau,t} \psi_\tau \, d\mu_\tau,$$

where

$$G_t = \bigoplus_{k=0}^{\infty} L^k m_t^k, \qquad G_{\tau,t} = \bigoplus L^k m_{\tau,t}^k,$$

and the functions $G_t \varphi_0$ and $\psi_{\tau,t} = G_{\tau,t} \psi_\tau$ are solutions of the homogeneous equations

$$\varphi_t = \varphi_0 \oplus \int_{(0,t]}^{\oplus} L \varphi_\tau \, d\mu_\tau, \qquad \varphi|_{t=0} = \varphi_0$$

and

$$\psi_{\tau,t} = \psi_\tau \oplus \int_{[\tau,t]}^{\oplus} L \psi_{s,t} \, ds, \qquad \psi|_{t=\tau} = \psi_\tau,$$

respectively.

Let us say a few words about stationary solutions. Obviously, each solution of Eq. (1.25) is a decreasing function of t. Suppose that the function $f(\tau)$ determining the measure μ_τ is nondecreasing and that the limit $\lim_{\tau \to \infty} f(\tau)$ exists and is equal to f_∞. In this case, if the limit $\varphi_\infty = \lim_{t \to \infty} \varphi_t$ exists, then

$$\int_{(0,\infty)}^{\oplus} \varphi_\tau(x) \, d\mu_\tau = \varphi_\infty(x-) \odot f_\infty,$$

and consequently, φ_∞ is a solution of the stationary equation

$$\varphi(x) = \varphi_0(x) \oplus \psi(x) \oplus L\varphi(x) \odot f_\infty,$$

where

$$\psi = \int_{(0,\infty)}^{\oplus} \psi_\tau \, d\mu_\tau.$$

Note that sometimes it is useful to consider infinite idempotent measures, that is, additive set functions on X defined only on subsets with compact closure. Such measures correspond to functionals in $(C_0(X))^*$, and the results of this section and of the next section are valid for these measures with obvious modifications.

The inner product

$$\langle f, h\rangle_A = \inf_x f(x) \odot h(x) = m_f(h) \tag{1.26}$$

is an idempotent analog of the conventional L^2 inner product of real functions. In different notation,

$$\langle f, h\rangle_A = \int_X^{\oplus} f(x) \odot h(x)\, d\mu_{\mathbb{1}}(x) = \int_X^{\oplus} h(x)\, d\mu_f(x).$$

Definition 1.2 1. A sequence of functions $f_n\colon X \to A$ is said to be *weakly convergent* to a function $f\colon X \to A$ on some semimodule of A-valued functions on X if for any function $h\colon X \to A$ in this semimodule we have

$$\lim_{n\to\infty} \langle f_n, h\rangle_A = \langle f, h\rangle_A. \tag{1.27}$$

2. A sequence of function $f_n\colon X \to A$ is said to be *convergent in measure* to a function $f\colon X \to A$ on a class of subsets of X if for each subset Ω from this class we have

$$\lim_{n\to\infty} \mu_{f_n}(\Omega) = \mu_f(\Omega), \tag{1.28}$$

where μ_f is the idempotent measure corresponding to f, i.e.,

$$\mu_f(\Omega) = \inf\{f(x) : x \in \Omega\}.$$

Theorem 1.11 *A sequence $\{f_n(x)\}$ of functions bounded below is weakly convergent to $f(x)$ on the set of upper semicontinuous functions tending to $\mathbb{0} = +\infty$ at infinity if and only if the sequence is bounded below and is convergent to $f(x)$ in measure on open subsets with compact closure.*

Proof. Suppose that the convergence (1.27) holds for upper semicontinuous functions $h(x)$ tending to $\mathbb{0}$ at infinity. Let Ω be an open subset with compact closure. We take $h(x)$ equal to the indicator function

$$\chi_\Omega(x) = \begin{cases} \mathbb{1} & \text{for } x \in \Omega, \\ \mathbb{0} & \text{for } x \notin \Omega, \end{cases} \tag{1.29}$$

which obviously belongs to the cited class, and obtain the limit equation (1.28) for Ω.

Conversely, suppose that the convergence (1.28) holds for open subsets Ω with compact closure. Then the convergence (1.27) holds for functions $h(x)$ of the form (1.29) and for their linear combinations (in the sense of the operations $\oplus$ and $\odot$). Let c be a constant such that $c \le f(x)$ and $c \le f_n(x)$ for all n and x. Let h be upper semicontinuous and tend to $\mathbb{0}$ at infinity. Then each set $\Omega_a = \{x : h(x) < a\}$ is open and has a compact closure for $a \ne \mathbb{0}$. Let $h_0 = \inf h(x)$. Let us introduce a finite partition $h_0 = a_0 < a_1 < \cdots < a_n = \mathbb{0}$ of the interval $[h_0, \mathbb{0}]$ so that the points $a_i + c$ form an ε-net on the set $[h_0 + c, \mathbb{0}]$ in the sense of the metric ρ. As was mentioned above, the convergence (1.27) holds for the function

$$h_\varepsilon(x) = \bigoplus_{i=1}^{n-1} a_i \odot \chi_{\Omega_{a_i}};$$

moreover, we have

$$\rho(\langle f_n, h_\varepsilon \rangle_A, \langle f_n, h \rangle_A) < \varepsilon$$

for all $n \in \mathbb{N}$ and

$$\rho(\langle f, h_\varepsilon \rangle_A, \langle f, h \rangle_A) < \varepsilon.$$

Thus, by passing to the limit as $\varepsilon \to 0$, we obtain Eq. (1.27) for the function h, as desired.

Let us give three examples so as to illustrate the notions introduced.

Example 1.9 For any real function $a(x)$ continuous on the interval $[0, 1]$, the sequence $f_n(x) = a(x) \cos nx$ is weakly convergent on $C_0^\infty([0, 1])$, as well as convergent in measure on open sets, to the function $f(x) = -|a(x)|$, which is the lower semicontinuous lower evolute of the sequence $f_n(x)$.

Example 1.10 Although the bounded sequence of continuous functions

$$f_n(x) = \begin{cases} 0 & \text{for } x \le 0, \\ nx & \text{for } x \in [0, 1/n], \\ 1 & \text{for } x \ge 1/n \end{cases}$$

on the real axis is monotone increasing and is pointwise convergent and weakly convergent on $C_0^\infty(\mathbb{R})$ to the lower semicontinuous step function

$$f(x) = \begin{cases} 0 & \text{for } x \le 0, \\ 1 & \text{for } x > 0, \end{cases}$$

this sequence is not convergent in measure on open sets with compact closure. Indeed, if this sequence were convergent to some function $f(x)$, then Eq. (1.22) for the intervals $\Omega = (\varepsilon_1, \varepsilon_2)$, $\varepsilon_1 > 0$, would imply that $f(x) \ge 1$ for $x \ge 0$, which contradicts (1.22) for the interval $\Omega = (0, 1)$.

Example 1.11 Although the sequence

$$f_n(x) = \begin{cases} 0 & \text{for } x \ne n, \\ -n & \text{for } x = n \end{cases}$$

is convergent to $f(x) = 0$ in measure on open sets with compact closure, it is not weakly convergent on upper semicontinuous functions tending to $\mathbb{0}$ at infinity.

Remark 1.4 Weak convergence on the set of all upper semicontinuous functions bounded below is equivalent to convergence in measure on all open sets, whereas weak convergence on the set of upper semicontinuous functions equal to $\mathbb{0}$ outside some compact set is equivalent to convergence in measure on open sets with compact closure.

Let us now give a criterion for weak convergence on $C_{\mathbb{0}}^*(X)$.

Lemma 1.7 *Let f_n be weakly convergent to f in $C_{\mathbb{0}}^*(X)$. Then for any open set U we have $\overline{\lim}_{n\to\infty} \mu_{f_n}(U) \le \mu_f(U)$.*

Proof. By definition, $\mu_f(U) = \inf_{x\in U} f(x)$. Therefore, for any $\varepsilon > 0$ there exists an $x_0 \in U$ such that $f(x_0) < \mu_f(U) + \varepsilon$. Let us choose an $h \in C_{\mathbb{0}}(X)$ such that $\operatorname{supp}_{\mathbb{0}} h \subset U$, $h(x_0) = \mathbb{1} = 0$, and $h(x) \ge \mathbb{1}$ everywhere on X. Then $m_f(h) \ge f(x_0) < \mu_f(U) + \varepsilon$, and consequently, for sufficiently large n we have

$$\mu_{f_n}(U) \le m_{f_n}(h) < m_f(h) + \varepsilon < \mu_f(U) + 2\varepsilon$$

by virtue of the formula $\lim_{n\to\infty} m_{f_n}(h) = m_f(h)$. The lemma is proved.

Theorem 1.12 *Let $f\colon X \to A$ be a lower semicontinuous function. A sequence f_n is weakly convergent to f in $C_{\mathbb{0}}^*(X)$ if and only if for any $\varepsilon > 0$ and $x \in X$ there exists a neighborhood U of x such that*

$$\rho(\mu_{f_n}(V), \mu_f(V)) < \varepsilon$$

for any open set V satisfying $x \in V \subset \overline{V} \subset U$ and for $n > N$, where N depends on V.

Proof. Necessity. Let us take an arbitrary $x \in X$. We carry out the argument for the case $f(x) \ne \mathbb{0}$ (the case $f(x) = \mathbb{0}$ is similar). Since f is semicontinuous, it follows that for any $\varepsilon > 0$ there exists a neighborhood U of the point x with compact closure $\overline{U}$ and a continuous function $h \in C_{\mathbb{0}}(X)$ such that $h = \mathbb{1} = 0$ in $\overline{U}$, $h \ge \mathbb{1}$ everywhere, and $f(x) \ge m_f(h) > f(x) - \varepsilon$. Then for any open set V with $x \in V \subset \overline{V} \subset U$ we have $f(x) \ge \mu_f(V) \ge m_f(h) > f(x) - \varepsilon$. Since $m_{f_n}(h) \to m_f(h)$ as $n \to \infty$, it follows that for large n we have

$$\mu_{f_n}(V) \ge m_{f_n}(h) \ge m_f(h) - \varepsilon > f(x) - 2\varepsilon \ge \mu_f(V) - 2\varepsilon.$$

In conjunction with Lemma 1.7, this inequality obviously implies the necessity of the condition given in the theorem.

Sufficiency. Let $h \in C_{\mathbb{0}}(X)$ and $\operatorname{supp}_{\mathbb{0}} h \subset K$, where K is a compact subset of X. We must prove that $m_{f_n}(h) \to m_f(h)$ as $n \to \infty$. Choose some $\varepsilon > 0$. Since h is continuous on the compact set K, it is uniformly continuous.

Hence, there exists a neighborhood D of the diagonal in $X \times X$ such that $\rho(h(x), h(y)) < \varepsilon$ for any $(x, y) \in D$. Furthermore, for each $x \in K$ let us construct a sufficiently small neighborhood V_x such that $\rho(\mu_{f_n}(V_x), \mu_f(V_x)) < \varepsilon$ for large n and $V_x \times V_x \subset D$. Let us choose a finite cover $\{V_{x_j}\}_{j=1,\dots,k}$ of K. Then $\inf_{x\in V_j}(f_n + h)$ is close to $\inf_{x\in V_j}(f + h)$ in the metric ρ on A for each j and for sufficiently large n, and consequently, $m_{f_n}(h)$ is close to $m_f(h)$. The theorem is proved.

Let us now establish an idempotent analog of the classical Banach–Alaoglu theorem. We omit the proof, which follows the classical counterpart and uses Tichonoff's theorem on products of compact sets.

Theorem 1.13 *The unit ball in the metric space $(C_{\mathbb{0}}^\infty(X))^*$ is compact in the weak topology; that is, any uniformly bounded sequence $f_n : X \to A$ contains a subsequence weakly convergent on $C_{\mathbb{0}}^\infty(X)$.*

Corollary 1.2 *Let $f_n : X \to A$ be a sequence such that $\lim_{n\to\infty} m_{f_n}(h)$ exists for every $h \in C_{\mathbb{0}}^\infty(X)$. Then the f_n are uniformly bounded ($f_n(x) \geq M$ for all x and n) and weakly converge on $C_{\mathbb{0}}^\infty(X)$ to their lower semicontinuous evolute.*

This corollary can also be proved directly or derived from the following general statement, whose detailed proof can be found in [**?**].

Theorem 1.14 *For any $h \in C_{\mathbb{0}}(X)$ and any sequence $f_n : X \to A$, we have*

$$\overline{\lim_{n\to\infty}}\, m_{f_n}(h) = m_f(h),$$

where f is the lower semicontinuous evolute of the sequence f_n.

In closing the section, let us note that the δ-functions in idempotent analysis are given by the indicator functions of points:

$$\delta_y(x) = \begin{cases} \mathbb{1} & \text{for } x = y, \\ \mathbb{0} & \text{for } x \neq y. \end{cases}$$

Indeed,

$$m_{\delta_y}(x) = \inf(\delta_y(x) \odot h(x)) = h(y).$$

For $X = \mathbb{R}^n$, simple delta-shaped sequences can be constructed of smooth convex functions; for example, $\delta_y(x)$ is the weak limit of the sequence $f_n(x) = n(x-y)^2$. Thus, by virtue of the preceding, each linear functional (or operator) on $C_{\mathbb{0}}^\infty(\mathbb{R}^n)$ is uniquely determined by its values on smooth convex functions.

1.5. The Fourier–Legendre Transformation

The well-known Legendre transformation is an analog of the Fourier transformation in idempotent analysis of A-valued functions on $\mathbb{R}^n$.

Indeed, the general Fourier transformation takes complex-valued functions on a commutative locally compact group G to functions on the dual group $\widehat{G}$ according to the formula

$$(Fh)(\chi) = \int_G \chi(x)h(x)\,dx,$$

where $\chi \in \widehat{G}$ is a character of G, that is, a continuous homomorphism of G into the unit circle S^1 considered as the multiplicative subgroup of unimodular numbers in $\mathbb{C}$.

In idempotent analysis, the characters of G can naturally be understood as the homomorphisms of G into the multiplicative group of the number semiring; then, for $G = \mathbb{R}^n$, the set of characters is the set of usual linear functionals on $\mathbb{R}^n$, naturally identified with $\mathbb{R}^n$. Next, we replace the integral by inf and the usual multiplication by the multiplication $\odot = +$ and obtain the following formula for the Fourier transform of an A-valued function h on $\mathbb{R}^n$:

$$(Fh)(p) = \inf_x (px + h(x)). \tag{1.30}$$

We see that $(Fh)(-p)$ is the usual Legendre transform with the opposite sign.

The usual Fourier transformation satisfies a commutation formula with convolution and is an eigenoperator of the translation operator. The same properties are valid for the Legendre–Fourier operator. Let us state the corresponding results; we leave the elementary proof of these statements to the reader as an exercise.

Theorem 1.15 1. *The Fourier–Legendre transformation* (1.30) *takes the convolution* (1.5) *to the multiplication:*

$$F(\varphi \circledast \psi) = F\varphi \odot F\psi.$$

2. *Let T_a be the transformation operator acting on the functions $h\colon \mathbb{R}^n \to A$ according to the formula*

$$(T_a h)(x) = h(x + a).$$

Then

$$(FT_a)(p) = pa \odot F(p). \tag{1.31}$$

3. *The Legendre transformation is the value at $t = \pi/2$ of the resolving operator O_t of the Cauchy problem for the Hamilton–Jacobi equation for the oscillator,*

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left[x^2 + \left(\frac{\partial S}{\partial x}\right)^2\right] = 0.$$

For each t, the operator O_t is a linear operator with the integral kernel

$$f_t(x,y) = \frac{1}{2}\csc t \cdot (\cos t\,(x^2+y^2) - 2xy).$$

The operator O_t satisfies the relation

$$(O_t R_a^t h)(p) = ax \csc t \odot (O_t h)(p),$$

where

$$(R_a^t h)(x) = h(x+a) + ax\cot t + \frac{1}{2}\,a^2\cot t,$$

which generalizes Eq. (1.31) *and coincides with Eq.* (1.31) *for* $t = \pi/2$. *Thus,* O_t *is an eigenoperator of the invertible operator* R_a^t.