The system jumps are defined as follows. We are given nonnegative functions $p^j(x)$, $j = 1,\ldots,m$, and m transformations of the state space of the system $\Gamma^j : \mathbb{R}^N \to \mathbb{R}^N$, $j = 1,\ldots,m$. The probability density of a jump of the j^{th} type on a small time interval $[t, t+dt)$ equals $\alpha^j(t)p^j(x(t))$.

A *strategy* is defined as a sequence of controls

$$\alpha_1(s|t,x), \alpha_2(s|t_1,x_1,t,x), \ldots, \alpha_{n+1}(s|t_n,x_n,t_{n-1},x_{n-1},\ldots,t,x),$$

where t_n is the *time* of the n^{th} jump and x_n is a *state* of the system immediately *after* the n^{th} jump.

It is required to choose a strategy β to maximize (minimize) the expected value of an additive functional on a fixed time interval, *viz.*

$$\sup_\beta E_x^\beta \sum_{n=1}^{\infty} q(x_n, t_n). \tag{1.27}$$

Besides the basic scheme problems, examples of this type of problem are given by some models of scientific technical progress considered in Arkin *et al.* (1976). The simplest of these has the following formulation. Assume that the state of the economy is described by a two dimensional vector $(x(t), y(t))$, where the first coordinate describes the output of the (industrial) *production* sector and the second that of the *science* sector. At each moment of continuous time, the resources $x(t)$ produced by the production sector are divided between the two sectors. Let $\alpha(t)$ be the fraction of production directed to industry, $0 \le \alpha \le 1$. The science sector develops production methods which define the productivity of the industrial sector, and the discovery of new methods is stochastic with the distribution of discovery epochs depending on the amount of resources invested in science. In the model, *productivity* is reflected by the coefficient k appearing in the differential equation describing the motion of the industrial sector, $\dot{x}(t) = k\alpha(t)x(t)$, which is changed by a jump at a random time τ, and the local probability of a jump in an interval $[t, t+dt)$ equals $y(t)dt$. So the system (1.26) for this model has the form ($x =: x_1$, $y =: x_2$, $\alpha =: \alpha^1$, $1-\alpha =: \alpha^2$),

$$\dot{x} = k\alpha x, \qquad \dot{y} = (1-\alpha)x,$$

with $m = 1$ and the function $p^1(x,y) := y$. A difference from the general case arises from the fact that instead of changes in the state of the system by jumps, changes in the form of (1.26) describing the system evolution occur. However, it is not difficult to construct a model permitting jumps of both types.

In Arkin *et al.* (1976) the optimal synthesis is given for the problem of minimizing the expected time to achieve a fixed level of production, assuming that only one jump is possible.

Returning to the consideration of the general case, as with the problems of the basic scheme, the problems considered may be divided into problems with a finite and an infinite number of jumps. In both cases, it may in principle be assumed that it is sufficient to consider the problem with one jump. In the first case, after a jump we have the known function $q(s,x)$, the profit function; in the second case, we may take as the profit function a function which coincides with the value function (plus the known function). The following two questions are of interest:

1. What conditions on the functions $q(s,x)$, $p^j(x)$ and the transformation Γ^j are required to ensure a smooth value function?
2. In the problem with an infinite number of jumps, a synthesis satisfying the Pontryagin maximum principle may be constructed under the assumption that the value function $F(s,x)$ is smooth. If the synthesis obtained is regular (see Boltyanski 1969), then is the function $F(s,x)$ actually smooth and is this synthesis optimal?

1.11 Results obtained

One of the main questions considered in this book concerns the behaviour of the value function $V_\nu(\xi)$ for an arbitrary hypothesis matrix $\{\lambda_i^j\}$ on a large time interval, where $V_\nu(\xi) := \sup_\pi V_\nu^\pi(\xi)$, and $V_\nu^\pi(\xi)$ is the expected number of jumps in the interval to time ν using strategy π in the basic scheme.

The first result in this direction was obtained in the previously quoted work by Feldman (1962). In particular, it was shown there that in the symmetric case with $m = N = 2$ (see the end of §1.2) the function $W_\nu(\xi) := \nu\lambda^2 - V_\nu(\xi)$ has a finite limit as $\nu \to \infty$. For an *arbitrary* hypothesis matrix the value function $V_\nu(\xi)$ is bounded above by the value $\nu\sum_{i=1}^N \xi_i\lambda_i$, where $\lambda_i := \max_{1\le j\le m}\lambda_i^j$. The expression $\nu\sum_{i=1}^N \xi_i\lambda_i$ has a clear interpretation. It represents the maximum value of profit that can be obtained by the statistician if when he chooses a strategy he knows which of the hypotheses H_i have been realized. (Formally, the inequality $V_\nu(\xi) \le \nu\sum_{i=1}^N \xi_i\lambda_i$ is easy to obtain, for example, from the convexity of the value function $V_\nu(\xi)$.) Therefore, it is natural to call the difference $\nu\sum_{i=1}^N \xi_i\lambda_i - V_\nu^\pi(\xi)$ the *loss* associated with strategy π. (For an optimal strategy this quantity is also often termed the *expected value of perfect information* (*EVPI*).) Denote it by $W_\nu^\pi(\xi)$, in both the discrete and continuous time cases. The value of loss for a strategy which minimizes the loss or, equivalently, maximizes the function $V_\nu^\pi(\xi)$, will be called the *loss function* $W_\nu(\xi)$.

Thus we have in the symmetric 2×2 case that $\lim_{\nu\to\infty} W_\nu(\xi) := W(\xi) < \infty$. Moreover, $W(\xi)$ coincides with the loss function on the infinite time interval and it may be shown that this agreement holds for an *arbitrary* $m \times N$ hypothesis matrix.

It turns out that all hypothesis matrices $\{\lambda_i^j\}$ may be divided into *two* classes. For all matrices in the first class

$$\lim_{\nu\to\infty} W_\nu(\xi) = W(\xi) < \infty, \tag{1.28}$$

while for those in the second class

$$\lim_{\nu\to\infty} W_\nu(\xi) = \infty\,. \tag{1.29}$$

The precise difference between these two classes will be explained in §3.1. Here it suffices to mention that the first matrix class contains all matrices which have no coinciding elements in their columns (in particular the symmetric 2×2 case, considered by Feldman (1962)). The only 2×2 matrices in the second class are matrices of type $\lambda_1^1 := a$, $\lambda_1^2 := \lambda_2^2 := b$, $\lambda_2^1 = c$, $0 \le a < b < c < 1$. This is the case of the *one-armed bandit* considered by Bellman (1956). Therefore, matrices of the first class will be called *F-matrices*, and those of the second class *B-matrices*.

The theorems establishing (1.28) and (1.29) for discrete time are proved in §§3.2 and 3.3. In §3.4 the optimal strategies for a 2×2 matrix are described, and at the end of the third chapter the scheme with resource sharing mentioned above is analysed.

In §§5.2 and 5.3 the optimal synthesis for an arbitrary 2×2 hypothesis matrix in continuous time is derived for the problem of loss minimization on finite and infinite time intervals. Let λ_i^j, $i = 1,2$, $j = 1,2$ be the coefficients of an *arbitrary* 2×2 matrix and define

$$\delta^j := \lambda_1^j - \lambda_2^j, \qquad \varepsilon_i := \lambda_i^1 - \lambda_i^2, \qquad \varepsilon := \varepsilon_1 - \varepsilon_2 = \delta^1 - \delta^2.$$

Without loss of generality (see §5.2), all matrices can be divided into the following *five* classes:

O) $\varepsilon_1 \cdot \varepsilon_2 \ge 0$

A) $\varepsilon_1 < 0, \quad \varepsilon_2 > 0, \quad \delta^1 < \delta^2 < 0$

B) $\varepsilon_1 < 0, \quad \varepsilon_2 > 0, \quad \delta^1 < \delta^2 = 0$

C) $\varepsilon_1 < 0, \quad \varepsilon_2 > 0, \quad 0 < \delta^2 < -\delta^1$

D) $\varepsilon_1 < 0, \quad \varepsilon_2 > 0, \quad 0 < \delta^2 = -\delta^1$.

For an infinite time interval the optimal synthesis is described by Theorem 5.2. It is more convenient to present this result in the variables $\eta := \tilde\eta(\xi) := \ln(c\xi/(1-\xi))$ where $\xi := \xi_1$, $c := -\varepsilon_1/\varepsilon_2$ (*cf.* (1.2)). It is proven that, with the exception of the trivial Case O, when the losses are equal to 0, and in Case B, when the losses are infinite, the half plane $\{(t,\eta) : t \ge 0\}$ is divided in two parts by the straight line $\eta = 0$. The optimal control depends only on the values of process $\eta(t) := \tilde\eta(\xi(t))$ and is given by the following synthesis: for $\eta < 0$, the control $\alpha = 1$ is optimal and for $\eta > 0$, the control $\alpha = 0$ is optimal.

The optimal trajectories of process $\eta(t) = \tilde\eta(\xi(t))$ in the intervals between jumps behave in the following way. In Cases C and D, the motion in the (t,η) plane is given by the differential equation $\dot\eta = -[\delta^1\alpha + \delta^2(1-\alpha)]$, which with $\alpha := 0$ or $\alpha := 1$ defines an angled straight line trajectory pointed towards the line $\eta = 0$, followed by motion along this line (see Figure 1). The line $\eta = 0$ is a particular solution for the differential equation above, and from a neighbourhood of this curve (for an appropriate control value) the system always comes to it and subsequently moves along it. Such a special curve is called a *turnpike*. The turnpike has a special significance in the Pontryagin theory of optimal control (see Chapter 6).

In Case A, if $\eta(t) < 0$ motion occurs along an angled straight line towards the line $\eta = 0$ with control $\alpha = 1$. When the line $\eta = 0$ is reached the control is switched to $\alpha = 0$ and further movement is along another straight line in the half plane $\eta > 0$ moving away from line $\eta = 0$. However, all jumps are in a downward direction.

In all cases the deterministic movement described is interrupted by jumps of value $\gamma^1 := \ln(\lambda_1^1/\lambda_2^1)$ with control $\alpha := 1$ and $\gamma^2 := \ln(\lambda_1^2/\lambda_2^2)$ with control $\alpha := 0$ (see (1.13)). The jump direction is always *opposite* to the direction of motion along the straight lines. The intensity of jumps equals $p^1(\tilde{\xi}(\eta(t)))$ with control $\alpha = 1$ and $p^2(\tilde{\xi}(\eta(t)))$ with control $\alpha = 0$, where $\tilde{\xi}(\eta) := e^{\eta}/(c + e^{\eta})$ is the inverse function to $\tilde{\eta}(\xi)$. On the turnpike the jump saltus may be equal to either γ^1 or γ^2 (see Figure 1).

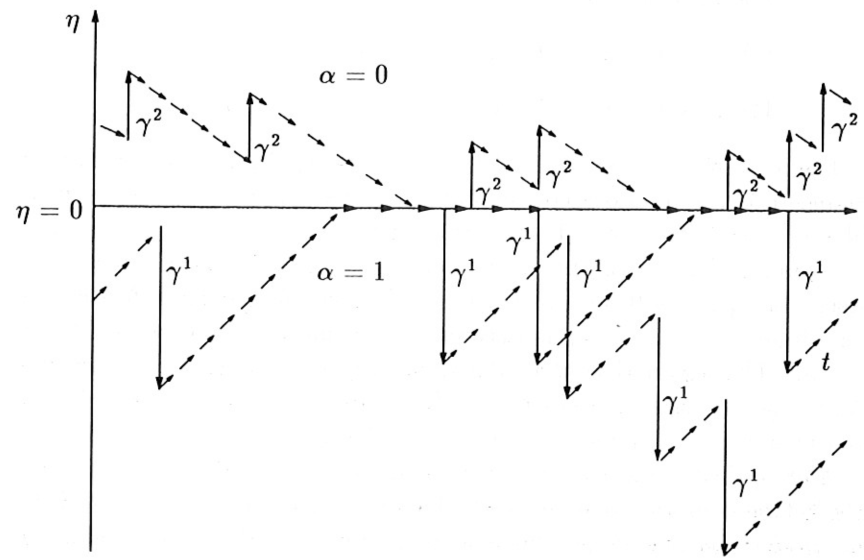

Figure 1
Optimal synthesis and optimal trajectories of the process $\eta(t)$ for the problem of loss minimization in continuous time over an infinite horizon with $m = N = 2$, Cases C and D.

Dotted arrows denote trajectories of the *a posteriori* probability of the first hypothesis on the time intervals between the jumps, and solid arrows denote the transformation of the *a posteriori* probability at jump moments corresponding to the optimal synthesis.

As well as describing the optimal synthesis, Theorem 5.2 also gives the values of the loss function $W(\xi)$ for all classes of hypothesis matrix.

In §5.3 the analogous problem for the case of a *finite* time interval is considered and a qualitative description of the optimal synthesis is given whose character stays the same as in the infinite horizon case.

According to Theorem 5.3, in Cases A–D for a finite horizon the halfplane (t, η), where t is the *time remaining*, is again divided in two parts, not by a straight line $\eta = 0$, but rather by a *curve* $\eta = l(t)$ with a kink at $t = t_*$ and the description of the behaviour of optimal trajectories remains the same (see Figure 2).

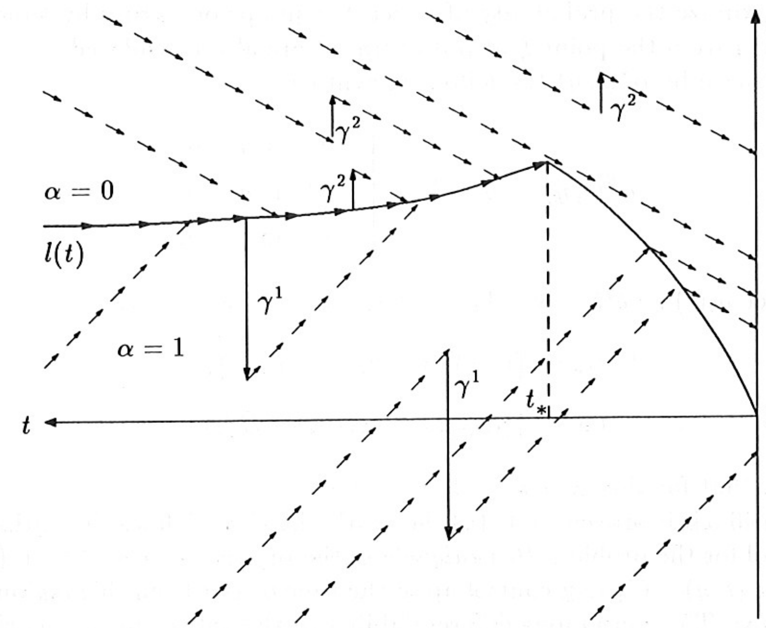

Figure 2
Optimal synthesis for the problem of loss minimization in continuous time with $m = N = 2$ (in coordinates (η, t), where t is time remaining), Case C.

In Cases A and B we have $t_* = \infty$ and the curve $\eta = l(t)$ is a switching-line, in Case D we have $t_* = 0$ and the curve $\eta = l(t)$ is a turnpike (i.e. a particular solution of the dynamics), while in Case C we have $0 < t_* < \infty$ and the part of the curve $\eta = l(t)$ for $t > t_*$ is the turnpike and the part for $0 < t < t_*$ is a switching-line. The curve $\eta = l(t)$ for time remaining $t = 0$ ends at the point $\eta = 0$ corresponding to the point ξ_*, which is the root of the equation

$$p^1(\xi) = p^2(\xi), \text{ i.e. } \xi\lambda_1^1 + (1-\xi)\lambda_2^1 = \xi\lambda_1^2 + (1-\xi)\lambda_2^2.$$

In Case C, whether the turnpike can be continued from t at t_* to infinity is an open question.

For the symmetric case, as well as the problem of maximization of the number of successes, the problems B_k in which it is necessary to maximize the probability of at least k jumps occurring by time ν starting from the point ξ at initial time t are also considered.

Denote by $\alpha^*(t,\eta)$ the following synthesis

$$\alpha^*(t,\eta) := \alpha^*(\eta) := \begin{cases} 0 & \text{if } \eta > 0 \\ 1/2 & \text{if } \eta = 0 \\ 1 & \text{if } \eta < 0, \end{cases}$$

and let $\eta_k(t) := \delta^2 t + (k-1)\gamma^1$, where $\gamma^1 := \ln(\lambda^1/\lambda^2)$, and

$$\overline{G}_k := \{(t,\eta) : t \geq 0, |\eta| < \eta_k(t)\},$$

$$G_k := \{(t,\eta) : t \geq 0, (t,\eta) \notin \overline{G}_k)\}.$$

(Note that for this case $\delta^2 > 0$, $\gamma^1 < 0$.)

In §6.5 it is proved that: (a) the synthesis $\alpha^*(\eta)$ defines the optimal control for the problem B_k (*uniquely* in the region $(\nu - t, \eta) \in G_k$); (b) for $(\nu - t, \eta) \in \overline{G}_k$ *any* control up to the time of exit from this region is optimal. The continuous differentiability of the value function is also proved and a simple recurrence formula is given for the value function in the area $\overline{G}_k$.

Figure 3 depicts (dotted lines) trajectories of the process $\eta(s)$ corresponding to the optimal control in the problem B_k. The small arrows represent trajectories of the *a posteriori* probabilities in the time intervals *between* jumps and the straight lines represent the boundaries of the areas G_i, $i = 1, \ldots, k$. The *indifference area* $\overline{G}_k$ is shaded.

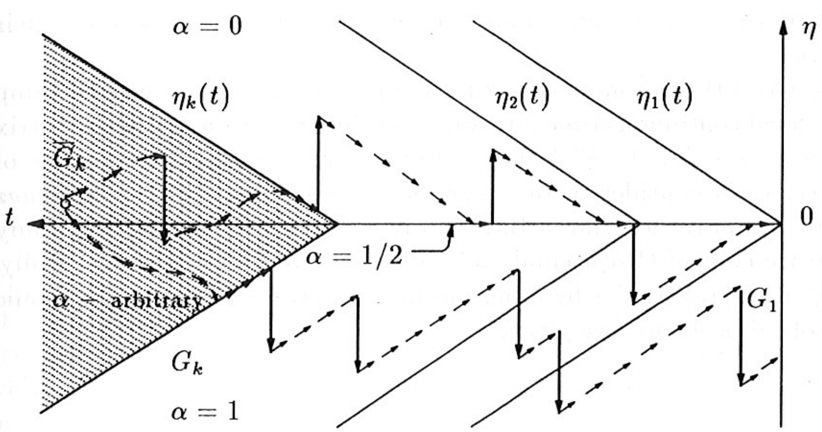

Figure 3
Optimal synthesis for the problem B_k in continuous time the symmetric case with $m = N = 2$ (in coordinates (η, t), where t is time remaining).

Since it is immediate that the synthesis $\alpha^*(\eta)$ determines the unique optimal control for *all* problems B_k, then from the equality $E\zeta = \sum_{k=1}^{\infty} P\{\tau_k \leq \nu\}$, where ζ is the number of jumps up to time ν and τ_k is the time of the k^{th} jump, we easily obtain that the strategy consisting of the control $\alpha^*(\eta)$ is the unique optimal strategy in the problem of maximizing the expected jump number (minimization of loss) in the symmetric case. This statement is analogous to the statement in the discrete time case and coincides with the corresponding results in §5.3. (The existence of an indifference area in the discrete time case was previously unmentioned.)

In Chapter 6 the approach of applying the Pontryagin maximum principle described above is studied. The results of this chapter mainly have the character of formulations and do not pretend to be complete.

In the final, seventh, chapter some special problems related to the basic scheme are considered. For a symmetric 2×2 hypothesis matrix and the case of discrete time, the problem of *best discrimination of* *hypotheses* is considered. The optimal strategy for this problem will *not* be the same for all symmetric matrices, but will depend on their entries.

In §7.3 the problem of maximizing the probability of the first jump in a fixed continuous time interval is considered for a hypothesis matrix of size $m \times N$. In §7.4 the problem of maximizing the number of successes is considered for a symmetric 2×2 matrix in a *minimax* formulation rather than a Bayesian one. In §7.5 the interesting study of price control (Rothschild 1974) discussed in §1.1 is given. Finally, in §7.6 several studies by other authors are presented briefly and some unsolved problems are given.

2 PROBLEM FORMULATION AND SOLUTION METHODS IN DISCRETE TIME

The precise formulation of the basic scheme as a problem of control of a stochastic process in discrete time is given in §2.1. In §2.2 the *general* problem of sequential control with incomplete information—of which the basic scheme can be considered a particular case—is given. In presenting this section we follow mainly the fundamental work of Schäl (1979) and the book of Dynkin & Yushkevitch (1976). In the following two sections some known general statements about the existence of optimal strategies, the relationship between the value function and the solution of the Bellman optimality equation and some other matters are presented. Related theorems in different forms are published in many articles and books and therefore as a rule we give these theorems in a form convenient for our presentation without precise reference. In §2.5 the equations governing the evolution of the *a posteriori* probabilities of hypotheses are given. We also introduce here the process $\eta(n)$ corresponding to the logarithmic likelihood function for these *a posteriori* probabilities and consider its separation into three processes possessing relatively simple properties. These basic results will be used in Chapter 3. The main contents of this section were originally presented in Presman & Sonin (1979).

2.1 Formulation of the basic scheme as a control problem

We introduce the following notation. We will always consider any vector as a *row-vector*. If y is a vector, then $y*$ denotes the corresponding column-vector and diag y denotes the diagonal matrix with the elements of the vector y on the diagonal. Let $S^k := \{y = (y_1, \ldots, y_k) : \sum_{i=1}^{k} y_i = 1, y_i \geq 0, i = 1, \ldots, k\}$ be a $(k-1)$-dimensional simplex; then $\widehat{S}^k$ denotes the set of its *extreme points*, i.e. $\widehat{S}^k :=$