3 SOLUTIONS OF SOME PROBLEMS IN THE BASIC DISCRETE TIME SCHEME

In this chapter we move from the study of general discrete time sequential control problems with incomplete information to some problems of the basic scheme.

Major attention is given in this chapter to a question which was formulated in Chapter 1. Namely, what is the asymptotic behaviour of the value function in the problem of maximizing the number of successes (minimization of loss) for an *arbitrary* hypothesis matrix $\Lambda = \{\lambda_i^j\}$ when the number of observations tends to infinity?

As mentioned in Chapter 1, all hypothesis matrices are divided into two classes and for the matrices in one of these classes—F-matrices—the loss to infinity is *finite*, while for the other class—B-matrices—it is *infinite*.

The corresponding statements are formulated in Theorem 3.1 (the proof of one of these statements will be deferred to §7.4). The proof is based on an investigation of the behaviour of some "reasonable" strategies and is a modification of that in Presman & Sonin (1979).

It seems that, with the exception of some particular cases, it is impossible in general to derive a description of optimal strategies in a sufficiently compact form. Descriptions of optimal strategies for finite and infinite observation horizons for the case of a 2×2 matrix are given respectively in Theorems 3.2 and 3.3. Our proofs are somewhat different from well-known proofs of these statements (see respectively Feldman 1962; De Groot 1970).

The chapter concludes with a section in which it is proved that in the scheme with sharable resources (see §1.6) the wider class of action rules relative to the basic scheme does not change the loss function. This fact is useful in the consideration of questions concerning the closeness of the discrete and continuous time basic schemes, since, as will be explained in §4.1, the continuous time basic scheme is a natural generalization of the scheme with sharable resources, rather than of the basic discrete time scheme.

3.1 F-matrices and B-matrices

In §2.4 it was shown that in solving problems of the basic scheme it is possible to replace a consideration of action rules by randomized Markov strategies, i.e. a sequence of functions $\pi := \{\pi_n(\xi), n = 1,2,\ldots\}$. For a fixed *a priori* distribution ξ^0, the action rule corresponding to a strategy π has the form $\beta := \{\pi_n(\xi(n)\}$, where the *a posteriori* probability $\xi(n)$ is given by formula (2.18) with $\xi := \xi^0$.

Let

$$\lambda_i := \max_{1\le j\le m} \lambda_i^j, \quad R_i := \{j : 1 \le j \le m,\, \lambda_i^j = \lambda_i\}. \tag{3.1}$$

It is convenient to speak of the *truth* of hypothesis H_i instead of about properties holding on the set $\{\theta_i = 1\}$. Accordingly, under the i^{th} hypothesis, R_i is the set of *best devices* (i.e. with maximum success probability) and λ_i is the common value of success probability for these devices.

In the case when for each fixed j all the λ_k^j are different, consider a strategy which prescribes the observation of a best device corresponding to an hypothesis with maximum *a posteriori* probability. We explain heuristically why such a strategy gives finite loss. Indeed, if the hypothesis H_i is true, then for finiteness of loss it is sufficient that after an (on average) *finite* time we will only use devices from the set R_i. But if for each fixed j, all the λ_k^i are different and hypothesis H_i is true, then (as was proved in §2.5) using *any* strategy the *a posteriori* probability of hypothesis H_i will quickly tend to 1, and for all other hypotheses to 0 (i.e. discrimination of hypotheses occurs). Therefore, applying the described strategy we will relatively quickly begin to use only devices from the set R_i.

The general situation is more complicated. Suppose that for some hypothesis H_i, a hypothesis H_k exists such that $\lambda_i^j = \lambda_k^j = \lambda_i$ for all $j \in R_i$ and $\lambda_i < \lambda_k \le 1$. In this case we will say that the hypothesis H_k is a *nuisance* for H_i and write $H_k > H_i$. If in this case the *a posteriori* probability of hypothesis H_i is close to 1, then under hypothesis H_k using the devices from R_i will on average increase the *a posteriori* probability of not only the hypothesis H_k but also of the hypothesis H_i (see formula (2.18)) and expectation of loss at each step in applying the strategy described above will be positive for *all* time.

22012 why H_k is a problem for H_i not vice versa

To have finite loss, it is necessary that with probability 1 we sufficiently quickly discriminate H_i from H_k. If H_k is a nuisance for H_i, we say the nuisance is *removable* if there exists a j such that either $\lambda_k^j = 1$ or $\lambda_k^j = 0$, $\lambda_i^j \ne 0$, and then we will call the j^{th} device *discriminating* for hypothesis H_i.

If in the presence of a removable nuisance we use the discriminating device, then for $\lambda_k^j = 1$ (correspondingly for $\lambda_k^j = 0$) in the case of failure (success) the *a posteriori* probability of the nuisance becomes 0 and with multiple use of the discriminating device such an event happens on average in *finite* time.

Therefore, for a strategy giving finite loss it is reasonable to use a device from the set R_i when the *a posteriori* probability of hypothesis H_i is greatest compared with all others and the *a posteriori* probability of all nuisance hypotheses with respect to H_i equals 0 (in this case the hypothesis H_i is called *preferable*).

Suppose all nuisances are removable and we utilize a stationary strategy which prescribes, in the presence of a preferable hypothesis, the observation of the devices which are the best according to this hypothesis with equal probability and, in the absence of a preferable hypothesis, the use of each device with equal probability. Then, as will be shown, the *a posteriori* probability of the true hypothesis will grow and we will relatively quickly come to the situation in which the loss is equal to 0.

Accordingly, a matrix Λ is called a *B-matrix* if its elements are such that there exists at least one nonremovable nuisance. All other matrices are called *F-matrices*. It is obvious that for $m = N = 2$ only the Bellman matrices (see §§1.11 and 2.5) are B-matrices.

The following theorem gives an answer to the question posed above about the finiteness of $W_\infty(\xi)$.

Theorem 3.1

(a) *For F-matrices, $W_\nu(\xi)\nearrow W_\infty(\xi) < \infty$ and there exists a stationary strategy realizing $W_\infty(\xi)$.*

(b) For B-matrices, $W_\nu(\xi)\nearrow+\infty$ and $\lim_{\nu\to\infty} W_\nu(\xi)/\nu = 0$. ■

We state an obvious corollary of this theorem.

Corollary 3.1 *For any matrix*

$$\lim_{\nu\to\infty} V_\nu(\xi)/\nu = \sum_{i=1}^{N} \xi_i\lambda_i. \tag{3.2}$$ ■

The convergence of $W_\nu(\xi)$ to $W_\infty(\xi)$ follows from Theorem 2.1. The statement about the existence of a stationary optimal strategy in the case in which for some ξ^0 the value of $W_\infty(\xi^0)$ is finite consists of Corollary 2.2. So, to prove (a) of the theorem it is only left for us to check that $W_\infty(\xi^0) < \infty$. Section 3.3 is devoted to this. In §3.4 it is shown that for B-matrices $W_\nu(\xi)\nearrow\infty$. The equality $\lim W_\nu(\xi)/\nu = 0$ will be proved only for $N = 2$ in §7.4, where it is also shown that for a minimax strategy the loss up to time ν has order $\sqrt{\nu}$. For the case $N > 2$, it is also not difficult to construct a strategy such that loss up to time ν will grow slower than ν, but we will not consider this question.

3.2 Loss for an F-matrix on an infinite time interval

In this section, we construct for an arbitrary F-matrix Λ a randomized stationary strategy $\overline{\pi}$ (given by a function $\overline{\pi}(\xi)$ on S^N) such that $W_\infty^{\overline{\pi}}(\xi^0) < \infty$ for any $\xi^0 \in S^N$.

To prove this property for $\overline{\pi}$, we will fix ξ^0 and estimate the right-hand side of the following equality, which holds by virtue of (2.10) and (2.3),

$$W_\infty^{\overline{\pi}}(\xi^0) = \sum_{i=1}^{N} \xi_i^0 \Big(\sum_{s=1}^{\infty} E_i^{\overline{\pi}} \sum_{j=1}^{m} (\lambda_i - \lambda_i^j)\overline{\pi}^j(\xi(s)) \Big), \tag{3.3}$$

where the coordinates of the vector $\xi(s)$ define the *a posteriori* probabilities of hypotheses at time s for the *a priori* distribution ξ^0.

First, let us introduce some notation. Let $A_i := \{k : H_k > H_i\}$ be the set of indices of nuisance hypotheses for the i^{th} hypothesis, $i = 1,\ldots,N$. For $0 < \varepsilon < 1$, let

$$\begin{aligned} D_i' &:= \{\xi : \xi_k/\xi_i < \varepsilon \text{ if } k \neq i; \xi_k = 0 \text{ if } k \in A_i\}, \\ D_0' &:= \{\xi : \xi \notin \bigcup_{i=1}^{N} D_i'\}. \end{aligned} \tag{3.4}$$

Obviously, the regions D_i' $(i = 1,\ldots,N)$ are mutually exclusive. Let us say that in region D_i' $(i = 1,\ldots,n)$, hypothesis H_i is *preferable*. This term indicates the fact that for $\xi \in D_i'$ the probability of any nuisance with respect to H_i is equal to 0 and the probabilities of all other hypotheses are small compared with the probability of hypothesis H_i.

Now we introduce formally the randomized stationary strategy $\overline{\pi}$ mentioned above and heuristically described in §3.1. This strategy has a simple and transparent probability interpretation (repeated here). If the *a posteriori* probabilities are such that there is a preferable hypothesis, then the strategy $\overline{\pi}$ prescribes the observation with equal probabilities of the devices which are best according to the preferable hypothesis. Otherwise, the strategy $\overline{\pi}$ prescribes the observation of *each* device with equal probability.

Formally, the strategy $\overline{\pi}$ is given by the following relations:

$$\begin{aligned} \overline{\pi} &:= \{\overline{\pi}_s(\xi),\ s = 1,2,\ldots\}, \\ \overline{\pi}_s(\xi) &:= \overline{\pi}(\xi) := (\overline{\pi}^1(\xi),\ldots,\overline{\pi}^m(\xi)), \\ \overline{\pi}^j(\xi) &:= \sum_{\{i: j\in R_i\}} \frac{1}{|R_i|} I\{D_i'\}, \quad j = 1,\ldots,m, \end{aligned} \tag{3.5}$$

where $R_0 := \{1,\ldots,N\}$, the R_i are defined in (3.1), and $|R|$ designates the number of elements in a finite set R.

If in each column of the matrix Λ all elements are different, then, according to (3.3), applying the strategy $\overline{\pi}$ the loss at time n is equal to 0 under hypothesis H_i only for $\xi(n) \in D_i'$. However, in general the loss will also be 0 when $\xi(n) \in D_l'$, where l is such that $\lambda_l = \lambda_i$ and $R_l \subseteq R_i$.

Correspondingly, for $i = 1,\ldots,N$, let

$$\overline{A}_i := \{l : \lambda_l = \lambda_i,\ R_l \subseteq R_i\}, \quad \overline{D}_i' := \bigcup_{l\in\overline{A}_i} D_l'. \tag{3.6}$$

From the definition of $\overline{\pi}(\xi)$ we have

$$\sum_{j=1}^{m}(\lambda_i - \lambda_i^j)\overline{\pi}^j(\xi) = 0 \quad \text{if } \xi \in \overline{D}_i'. \tag{3.7}$$

According to (3.3) and (3.7), $\overline{D}_i'$ is a region of zero loss when hypothesis H_i is true (i.e. on the set $\{\theta_i = 1\}$).

From (3.3) and (3.7), we may derive the estimate

$$W_\infty^{\overline{\pi}}(\xi^0) \le c \sum_{i=1}^{N} \xi_i^0 \sum_{s=1}^{\infty} P_i^{\overline{\pi}}\{\xi(s) \notin \overline{D}_i'\}, \tag{3.8}$$

where $c := \max_{i,j}\ (\lambda_i - \lambda_i^j)$. Therefore, to prove the finiteness of $W_\infty^{\overline{\pi}}(\xi^0)$ it is sufficient to show that if $\xi_i^0 > 0$ then

$$\sum_{s=1}^{\infty} P_i^{\overline{\pi}}\{\xi(s) \notin \overline{D}_i'\} < \infty. \tag{3.9}$$

The proof for each $i = 1, \dots, N$ is conducted similarly, and thus, without loss of generality, we will consider here *only the case* $i = N$, assuming $\xi_N^0 > 0$. Therefore, as in §2.5, all events are assumed conditional on the event $\{\theta_N = 1\}$ and in equalities between random values we shall correspondingly omit the multiplier in terms of θ_N. Also, for brevity, we shall replace $P_N^{\overline{\pi}}$ by P in the derivation.

As in §2.5, we first make a change of variables

$$\tilde{\eta}(\xi) := \left(\ln\frac{\xi_1}{\xi_N}, \dots, \ln\frac{\xi_{N-1}}{\xi_N}\right). \tag{3.10}$$

Such a change is appropriate, since on the set $\{\theta_N = 1\}$, according to (2.4) and (2.16)–(2.18), from $\xi_N^0 > 0$ it follows that $\xi_N(s) > 0$ for any s. Consider now the process

$$\eta(s) := \tilde{\eta}(\xi(s)). \tag{3.11}$$

By (2.93) and (2.94) the coordinates of the process $\eta(s)$ satisfy the relations

$$\eta(n) = \eta(n-1) + \sum_{j=1}^{m} a^j(n)[\gamma^{1j}\Delta X^j(n) + \gamma^{0j}(1 - \Delta X^j(n))]. \tag{3.12}$$

Images of the regions D_i' and $\overline{D}_i'$ under the mapping (3.10) we denote respectively by D_i and $\overline{D}_i$ and we will omit the index N in denoting D_N, $\overline{D}_N$, A_N, $\overline{A}_N$. So, for $\eta = (\eta_1, \dots, \eta_{N-1})$ and $a = -\ln\varepsilon$

$$D := D_N = \{\eta : \eta_k < -a,\ k \notin A,\ \eta_k = -\infty,\ k \in A\}$$

$$D_i = \begin{cases} \{\eta : \eta_i > a,\ \eta_k - \eta_i < -a & \text{for } k \ne i,\ k \notin A_i, \\ \qquad \eta_k = -\infty \text{ for } k \in A_i\} & \text{if } N \notin A_i \\ \varnothing & \text{if } N \in A_i \end{cases} \tag{3.13}$$

$$D_0 = \{\eta : \eta \notin \bigcup_{i=1}^{N} D_i\}.$$

Note that relation (3.13) for the case $N \in A_i$ is obtained from the condition $\xi_N(0) > 0$.

To prove (3.9) it is thus sufficient to show that

$$\sum_{s=1}^{\infty} P\{\eta(s) \notin \overline{D}\} < \infty, \tag{3.14}$$

where $\overline{D} = \bigcup_{l \in \overline{A}} D_l$. For reasons of simplicity, first consider the situation where for each $j = 1, \dots, m$ all λ_i^j are different from each other and different from 0 and 1. Then

$$\overline{D} = D = \{\eta : \eta_i < -a,\ i = 1, \dots, N-1\} \tag{3.15}$$

and (3.14) follows from the inequalities proved in Theorem 2.8, viz.

$$\sum_{s=1}^{\infty} P\{\eta_k(s) > -a\} < \infty, \quad k = 1, 2, \dots, N-1. \tag{3.16}$$

Relation (3.16) holds because all coordinates of the process $\eta(s)$ tend to $-\infty$ sufficiently quickly. In the general situation this is not so. With positive probability the process $\eta(s)$ may never enter regions where the true hypothesis is preferable.

We show below that hypotheses which do not belong to the set $\overline{A}$ will *not* be preferable a considerable part of the time. Moreover, if for a considerable share of a sufficiently long time interval some hypothesis with index k from the set $\overline{A}$ is preferable, then this hypothesis will stay preferable in the future with high probability.

To formulate these statements precisely, define for $k = 0, 1, \dots, N$ the events

$$\Omega_k^s := \{\omega : \sum_{r=1}^{s} I\{\eta(r) \in D_k\} \geq s/(N+1)\}. \tag{3.17}$$

The event Ω_k^s for $k \neq 0$ is defined by the condition that the k^{th} hypothesis is preferable at least $s/(N+1)$ times over the time interval from 1 to s and the event Ω_0^s is defined by the condition that there is no preferable hypothesis at least $s/(N+1)$ times. It is obvious that $\bigcup_{k=0}^{N} \Omega_k^s = \Omega$, and therefore to prove (3.14) it is sufficient to show that for all $k = 0, 1, \dots, N$

$$\sum_{s=1}^{\infty} P\left[\{\eta(s) \notin \overline{D}\} \cap \Omega_k^s\right] < \infty. \tag{3.18}$$

If k is such that $N \in A_k$, then from (3.13) $D_k = \varnothing$ and therefore

$$P\{\Omega_k^s\} = 0 \quad \text{for } N \in A_k. \tag{3.19}$$

So we will consider further only those k for which $N \notin A_k$.

We prove that for some $c > 0$, where c depends only on the initial point ξ^0, the following inequalities hold:

$$P[\{\eta(s) \notin D\} \cap \Omega_0^s] < c/s^2, \tag{3.20}$$

$$P\{\Omega_k^s\} < c/s^2 \quad \text{if } k \notin \overline{A}, \tag{3.21}$$

$$P[\{\eta(s) \notin D_k\} \cap \Omega_k^s] < c/s^2 \quad \text{if } k \in \overline{A}. \tag{3.22}$$

Since from $k \in \overline{A}$ we have $D_k \in \overline{D} := \bigcup_{l \in \overline{A}} D_l$, (3.18) follows from (3.20)–(3.22).

To prove the relations (3.20)–(3.22) we will consider the behaviour of the coordinates of the process $\eta(s)$.

First, we show that (3.20) is a consequence of the following inequalities for some α, $c > 0$ and any $s > 0$:

$$P[\{\eta_i(s) \neq -\infty\} \cap \Omega_0^s] \leq c/s^2 \quad \text{if } i \in A, \tag{3.23}$$

$$P[\{\eta_i(s) > -\alpha s\} \cap \Omega_0^s] \leq c/s^2 \quad \text{if } i \notin A. \tag{3.24}$$

Indeed, if s is sufficiently large, then $\alpha s > a$ and, from $\eta_i(s) < -\alpha s$ for all $i \notin A$, $\eta_i(s) = -\infty$ for all $i \in A$, it follows that $\eta(s) \in D$. Therefore for such s from (3.23) and (3.24) it follows that

$$P\left[\{\eta(s) \notin D\} \cap \Omega_o^s\right]$$
$$\leq P\left[\left[\left(\bigcup_{i \in A} \{\eta_i(s) \neq -\infty\}\right) \cup \left(\bigcup_{i \notin A} \{\eta_i(s) > -\alpha s\}\right)\right] \cap \Omega_0^s\right] \leq Nc/s^2.$$

For the remaining s, the inequality (3.20) may be assumed to hold by an appropriate choice of c. Thus, (3.20) follows from (3.23) and (3.24).

We show now that (3.21) follows from the fact that, for some α, $c' > 0$ and any $s > 0$, the inequality

$$P[\{\eta_k(s) > -\alpha s\} \cap \Omega_k^s] \leq c'/s^3 \quad \text{if } k \notin \overline{A} \tag{3.25}$$

holds.

Indeed, on the set Ω_k^s define the time τ_k^s as the (random) last exit time of the process $\eta(n)$ from the region D_k in the time interval $[1, s]$. From the definition of Ω_k^s it follows that

$$s/(N+1) \leq \tau_k^s \leq s, \quad \{\tau_k^s = r\} \cap \Omega_k^s \subseteq \{\tau_k^r = r\} \cap \Omega_k^r, \tag{3.26}$$

and from the definition of the region D_k it follows that for $\eta(s) \in D_k$, $\eta_k(s) > a$ holds. If s is sufficiently large, then $\alpha s/(N+1) > a$ and from (3.25) and (3.26) we obtain that for $k \notin \overline{A}$

$$\begin{aligned} P\{\Omega_k^s\} &= \sum_{s/(N+1) \leq r \leq s} P[\Omega_k^s \cap \{\tau_k^s = r\}] \\ &\leq \sum_{s/(N+1) \leq r \leq s} P[\Omega_k^r \cap \{\tau_k^r = r\}] \\ &\leq \sum_{s/(N+1) \leq r \leq s} P[\{\eta_k(r) > a\} \cap \Omega_k^r] \\ &\leq \sum_{s/(N+1) \leq r \leq s} P[\{\eta_k(r) > -\alpha r\} \cap \Omega_k^r] \\ &\leq \sum_{s/(N+1) \leq r \leq s} c'/r^3 < c/s^2, \end{aligned}$$

i.e. from (3.25) we have (3.21).

Finally, we define

$$\widehat{A}_k := \{l : \lambda_l = \lambda_k, R_k \subseteq R_l\} \cup A_k \tag{3.27}$$

and show that (3.22) is a consequence of the following inequalities for some $\alpha > 0$, $c' > 0$ and for any $s > 0$:

$$P[\{\eta_i(s) > -\alpha s\} \cap \Omega_k^s] \leq c'/s^3 \quad \text{if} \quad k \in \overline{A},\ i \notin \widehat{A}_k. \tag{3.28}$$

Indeed, from the definition of the strategy $\overline{\pi}$ and from (3.12), we obtain that for $\eta(r) \in D_k$, $k \in \overline{A}$, the equality $\eta_i(r+1) = \eta_i(r)$ holds for $i \in \widehat{A}_k$ with probability 1. The maximum possible positive jump of the other coordinates is bounded by some number which depends only on the elements of the matrix Λ. Therefore, for sufficiently large r it follows from $\eta(r) \in D_k$ and $\eta_i(r) > -\alpha r$ for all $i \notin \widehat{A}_k$ that $\eta(r+1) \in D_k$. Thus, using (3.26) we have

$$\begin{aligned}
&P[\{\eta(s) \notin D_k\} \cap \Omega_k^s] \\
&\quad = \sum_{s/(N+1) \leq r \leq s-1} P[\{\tau_k^s = r\} \cap \Omega_k^s] \\
&\quad \leq \sum_{s/(N+1) \leq r \leq s-1} P[\{\tau_k^r = r\} \cap \Omega_k^r] \\
&\quad \leq \sum_{s/(N+1) \leq r \leq s-1} P[\{\eta(r) \in D_k\} \cap \{\eta(r+1) \notin D_k\} \cap \Omega_k^r] \\
&\quad \leq \sum_{s/(N+1) \leq r \leq s-1} P[\{\cup_{i \notin \widehat{A}_k} \{\eta_i(r) > -\alpha r\} \cap \Omega_k^r] \\
&\quad \leq \sum_{s/(N+1) \leq r \leq s-1} c'/r^3 \leq c/s^2,
\end{aligned}$$

i.e. from (3.28) we have (3.22).

So, it remains to prove the inequalities (3.23)–(3.25) and (3.28). Consider the representation of process $\eta(n)$ as the sum of three processes introduced at the end of §2.5, viz.

$$\begin{aligned}
\eta(n) &= \eta^1(n) + \eta^2(n) + \eta^3(n) \\
\eta^1(0) &= \eta^3(0) = 0 \\
\eta^2(0) &= \eta(0),
\end{aligned} \tag{3.29}$$

where

$$\Delta\eta_k^1(n) := \sum_{j \in R_k''} a^j(n)(\gamma_k^{1j} - \gamma_k^{0j})\,(\Delta X^j(n) - a^j(n)\lambda_N^j) \tag{3.30}$$

$$\Delta\eta_k^2 := \sum_{j \in R_k''} a^j(n)\,(\gamma_k^{0j}(1 - \lambda_N^j) + \lambda_N^j\gamma_k^{1j}) \tag{3.31}$$

$$\Delta\eta_k^3 := \sum_{j \in R_k'} (a^j(n)\,(\gamma_k^{1j}\Delta X^j(n) + \gamma_k^{0j}(1 - \Delta X^j(n))). \tag{3.32}$$

Here $\gamma_k^{1j} := \ln(\lambda_k^j/\lambda_N^j)$, $\quad \gamma_k^{0j} := \ln[(1-\lambda_k^j)/(1-\lambda_N^j)]$,

$$R_k' := \{j : \lambda_k^j = 1,\ 0 \leq \lambda_N^j < 1 \ \text{ or } \ \lambda_k^j = 0,\ 0 < \lambda_N^j \leq 1\}, \tag{3.33}$$

$$R_k'' := \{j : j \notin R_k',\ \lambda_k^j \neq \lambda_N^j\}.$$

As already mentioned, the coordinates of the process $\{\eta^1(n), \mathcal{F}_n\}$ are martingales, and their increments are bounded by some fixed number. From this, by Lemma 2.5, for any $\alpha > 0$ there exists c such that for all $s > 0$ and $k = 1, \ldots, N-1$

$$P\{\eta_k^1(s) > \alpha s\} < c/s^3. \tag{3.34}$$

The increments of coordinates of process $\eta^3(n)$ are either equal to 0 or, if an increment is not 0, with positive probability less than some fixed $q < 1$ it takes some positive value (bounded by some constant) and with the complementary probability it takes the value $-\infty$. Therefore, for any α there exists a c such that for all $s > 0$ and $k = 1, \ldots, N-1$

$$P\{\eta_k^3(s) > \alpha s\} < q^s < c/s^3. \tag{3.35}$$

The process $\eta^2(n)$ is a nonincreasing process, since (see Lemma 2.4)

$$(1 - \lambda_N^j)\ln\frac{1-\lambda_k^j}{1-\lambda_N^j} + \lambda_N^j \ln\frac{\lambda_k^j}{\lambda_N^j} < -b < 0 \quad \text{if } \lambda_N^j \neq \lambda_k^j. \tag{3.36}$$

In general not all coordinates of process $\eta^2(n)$ are necessarily decreasing; however, we now show that with the occurrence of the appropriate event Ω_k^s the coordinates of the process $\eta^2(n)$ entering into the formulae (3.23)–(3.25) and (3.28) will go sufficiently far into the negative region with probability close to 1.

First we prove (3.25).

If $k \notin \overline{A}$ then there exists a $j \in R_k$ such that $\lambda_k^j = \lambda_k \neq \lambda_N^j$. From the definition of the strategy $\overline{\pi}$ and from (2.3) it follows that at those times r when hypothesis H_k is preferable the values $a^j(r)$ for this j will be independent Bernoulli random values with success probability $1/|R_k|$. There will be at least $s/(N+1)$ such times on the set Ω_k^s. From this it follows that for this j

$$P\left[\left\{\sum_{r=1}^{s} a^j(r) > t\right\} \cap \Omega_k^s\right] > 1 - F(1/|R_k|, s/(N+1), t), \tag{3.37}$$

where $F(p,n,t)$ is the value at the point t of the distribution function of a binomial random variable with probability of success p and number of tests n. From definition (3.31) and from (3.36), we find that for this k and j

$$P\left[\{\eta_k^2(s) < -\alpha s\} \cap \Omega_k^s\right] \geq P\left[\left\{\sum_{r=1}^{s} a^j(r) > \alpha s/b\right\} \cap \Omega_k^s\right].$$

From this, (3.37) and the exponential character of the tails of binomial distributions—which follows, for example, from Lemma 2.5—we obtain that if $k \notin \overline{A}$ then for some α, $c > 0$

$$P[\{\eta_k^2(s) > -\alpha s\} \cap \Omega_k^s] < c/s^3. \tag{3.38}$$

From (3.29), (3.34), (3.35) and (3.38) we have (3.25) and hence (3.21) also.

Next we prove (3.28).

If $k \in \overline{A}$, $i \notin \widehat{A}_k$, then there exists a $j \in R_k$ such that $\lambda_i^j \neq \lambda_k^j = \lambda_N$. For this j, inequality (3.36) is again true. Therefore if this j belongs to R_i'', then there exist α, $c > 0$ such that

$$P[\{\eta_i^2(s) > -\alpha s\} \cap \Omega_k^s] < c/s^3, \tag{3.39}$$

and if $j \in R_k'$, then

$$P[\{\eta_i^3(s) \neq -\infty\} \cap \Omega_k^s] < c/s^3. \tag{3.40}$$

From (3.34), (3.35), (3.39) and (3.40) we obtain (3.28).

Finally, we prove (3.24) and (3.23).

For any i there exists j such that $\lambda_i^j \neq \lambda_N^j$. As in the proof of (3.37), we obtain

$$P\left[\left\{\sum_{r=1}^{s} a^j(r) > t\right\} \cap \Omega_0^s\right] > 1 - F(1/m, s/(N+1), t), \tag{3.41}$$

and therefore, if this j belongs to R_i'', then there exist c and α such that

$$P[\{\eta_i^2(s) > -\alpha s\} \cap \Omega_0^s] < c/s^3, \tag{3.42}$$

and if $j \in R_i'$ (for $i \in A$ it may be assumed that this membership holds), then

$$P[\{\eta_k^3(s) \neq -\infty\} \cap \Omega_0^s] < c/s^3. \tag{3.43}$$

From (3.34), (3.35), (3.42) and (3.43) now follow (3.23) and (3.24).

This concludes the proof of the fact that

$$W_\infty^{\overline{\pi}}(\xi^0) < \infty. \tag{3.44}$$

3.3 Loss for a B-matrix on an infinite time interval

Let Λ be a B-matrix. Without loss of generality, re-indexing as necessary the hypotheses, it will be assumed that H_2 is a nonremovable nuisance for H_1, i.e.

$$\lambda_2^j = \lambda_1^j = \lambda_1 \ \text{ if } \ j \in R_1, \tag{3.45}$$

$$\lambda_2^j = 0 \ \text{ implies } \ \lambda_1^j = 0, \tag{3.46}$$

$$0 < \lambda_1 < \lambda_2 < 1. \tag{3.47}$$

We will show that for any action rule β and any ξ with $\xi_i > 0$ for $i = 1, 2$, $W_\infty^\beta(\xi) = \infty$. Suppose the contrary. If $W(\xi) < \infty$, then by (2.10) there exists an action rule $\overline{\beta}$ such that

$$\begin{gathered} \sum_{s=1}^{\infty} E_1^{\overline{\beta}}\Big(\sum_{j \notin R_1} a^j(s)\Big) < \infty, \\ \sum_{s=1}^{\infty} E_2^{\overline{\beta}}\Big(\sum_{j \notin R_2} a^j(s)\Big) < \infty. \end{gathered} \tag{3.48}$$

Since $a^j(s) \geq 0$, the expectations and sums in (3.48) may be interchanged.

The sets R_1 and R_2 are disjoint and $\sum_{j=1}^m a^j(s) = 1$, so that from (3.48) it follows that

$$E_1^{\bar\beta}\Big[\sum_{s=1}^{\infty}(1 - \sum_{j\in R_1} a^j(s))\Big] < \infty,$$
$$E_2^{\bar\beta}\Big[\sum_{s=1}^{\infty}\sum_{j\in R_1} a^j(s)\Big] < \infty. \tag{3.49}$$

Define the $\mathcal{F}_\infty$-measurable random variable $\tau \leq +\infty$ by

$$\tau := \inf\{s : \sum_{j\in R_1} a^j(r) = 1 \text{ for all } r > s\}. \tag{3.50}$$

From (3.49) we have that

$$P_1^{\bar\beta}\{\tau < \infty\} = P_1^{\bar\beta}\Big\{\sum_{s=1}^{\infty}(1 - \sum_{j\in R_1} a^j(s)) < \infty\Big\} = 1,$$
$$P_2^{\bar\beta}\{\tau < \infty\} \leq P_2^{\bar\beta}\Big\{\sum_{s=1}^{\infty}\sum_{j\in R_1} a^j(s) = \infty\Big\} = 0. \tag{3.51}$$

We show that under conditions (3.45)–(3.47) the two relations of (3.51) are contradictory.

Indeed, from (2.2)–(2.4) and (3.45) it follows that for any n and $n_1 > n$,

$$P_1^{\bar\beta}\Big\{\sum_{j\in R_1} a^j(s)=1,\ n+1 \leq s \leq n_1 | \mathcal{F}_n\Big\}$$
$$= P_2^{\bar\beta}\Big\{\sum_{j\in R_1} a^j(s) = 1,\ n+1 \leq s \leq n_1 | \mathcal{F}_n\Big\}$$

holds.

Taking limits as $n_1 \to \infty$, we obtain that this equality holds for $n_1 = \infty$, i.e.

$$P_1^{\bar\beta}\{\tau \leq n|\mathcal{F}_n\} = P_2^{\bar\beta}\{\tau \leq n|\mathcal{F}_n\}.$$

Therefore, by the finiteness of the σ-algebra $\mathcal{F}_n$ there exists a function $\gamma(\cdot)$ such that for $i = 1, 2$,

$$P_i^{\bar\beta}\{\tau \leq n|\mathcal{F}_n\} = \gamma(a(1), \Delta X(1), \ldots, a(n), \Delta X(n)).$$

From this we obtain that for $i = 1, 2$,

$$P_i^{\bar\beta}\{\tau \leq n\} = E_i^{\bar\beta}(P_i^{\bar\beta}\{\tau \leq n|\mathcal{F}_n\})$$
$$= E_i^{\bar\beta}(\gamma(a(1), \Delta X(1), \ldots, a(n), \Delta X(n))). \tag{3.52}$$

The number of possible trajectories $a(1), \Delta X(1), \ldots, a(n), \Delta X(n)$ are finite for fixed n and, by (3.46) and (3.47), if the probability of any of these trajectories is positive with respect to the measure $P_1^{\bar\beta}$, then it is positive with respect to the measure $P_2^{\bar\beta}$ also. Therefore, from (3.52) it follows that if

$$P_1^{\bar\beta}\{\tau \leq n\} > 0, \quad \text{then} \quad P_2^{\bar\beta}\{\tau \leq n\} > 0,$$

which contradicts (3.51).

So, in the case of a B-matrix, $W_\infty^\beta(\xi) = \infty$ for any action rule β and any ξ with $\xi_i > 0$ for $i = 1, \ldots, N$.

Remark 3.1 For a B-matrix of second order (i.e. the Bellman matrix) there does not exist a stationary asymptotic optimal strategy, since if at some point $\bar\beta^1(\xi) = 0$ and $\xi(0) = \xi$, then $\xi(n) \equiv \xi$ and $W_\nu^\beta(\xi) = \nu(1 - \xi_1)(\lambda_2 - \lambda_1)$, which contradicts (b) of Theorem 3.1. However, there do exist examples of B-matrices of large order which have such strategies. ■

3.4 Optimal strategies for the case $m = N = 2$

In the case $m = N = 2$, or as we will sometimes say, in the case of a 2×2 matrix, a description of optimal strategies in the problem of loss minimization (maximization of the number of successes) may be obtained in *explicit* form.

Let

$$\delta^j := \lambda_1^j - \lambda_2^j,\ \ \varepsilon_i := \lambda_i^1 - \lambda_i^2,\ \ \varepsilon := \varepsilon_1 - \varepsilon_2 = \delta^1 - \delta^2. \tag{3.53}$$

Since $\xi_1 + \xi_2 = 1$ it is convenient to consider a *scalar* value $\xi := \xi_1$ rather than a vector $\xi := (\xi_1, \xi_2)$. Similarly, we introduce $\beta(s) := \beta^1(s)$ and $\pi(s,\xi) := \pi^1(s,\xi)$ and define $W_s^\pi(\xi) := W_s^\pi(\xi, 1-\xi)$ and $W_s(\xi) := W_s(\xi, 1-\xi)$.

According to §2.4 the problem of loss minimization on the time interval $[0,\nu)$, $\nu \le \infty$, for the case $m = N = 2$ may be considered to be a homogeneous Markov model whose states are points $\xi \in [0,1]$ and for each state there exist two controls having values 1 and 2. Transition probabilities in this notation are given by the formulae

$$\begin{aligned} P\{\xi(n) &= \Gamma^{1j}\xi \mid \xi(n-1) = \xi,\ a(n) = j\} = p^j(\xi), \\ P\{\xi(n) &= \Gamma^{0j}\xi \mid \xi(n-1) = \xi,\ a(n) = j\} = 1 - p^j(\xi), \qquad (3.54) \\ &\qquad j = 1,2, \end{aligned}$$

where

$$p^j(\xi) := \lambda_1^j \xi + \lambda_2^j (1-\xi) = \xi\delta^j + \lambda_2^j$$

$$\Gamma^{1j}\xi := \begin{cases} \xi\lambda_1^j / p^j(\xi) & \text{if } p^j(\xi) \ne 0 \\ \xi & \text{if } p^j(\xi) = 0 \end{cases} \qquad (3.55)$$

$$\Gamma^{0j}\xi := \begin{cases} \xi(1-\lambda_1^j)/(1-p^j(\xi)) & \text{if } p^j(\xi) \ne 1 \\ \xi & \text{if } p^j(\xi) = 1. \end{cases}$$

The cost functions are given by

$$q^j(\xi) := \xi(\lambda_1 - \lambda_1^j) + (1-\xi)(\lambda_2 - \lambda_2^j),$$

$$\text{where} \qquad \lambda_i := \max_{i=1,2}(\lambda_i^1, \lambda_i^2), \qquad (3.56)$$

and the operators T^j have the form

$$\begin{aligned} T^j f(\xi) &:= q^j(\xi) + M^j f(\xi) \\ &:= q^j(\xi) + p^j(\xi) f(\Gamma^{1j}\xi) + (1 - p^j(\xi)) f(\Gamma^{0j}\xi). \qquad (3.57) \end{aligned}$$

Formulae (3.54)–(3.56) are a special case of the general formulae (2.82)–(2.85) for $m = N = 2$, $\xi_1 = \xi$, $\xi_2 = 1 - \xi$.

Changing as necessary the indices of devices and hypotheses, it may be assumed that $|\delta^1| \ge |\delta^2|$, $\delta^1 \le 0$. If $\varepsilon_1 \cdot \varepsilon_2 < 0$, then from this it is easily obtained that $\delta^1 \ne \delta^2$, $\delta^1 < 0$, $\varepsilon_1 < 0$, $\varepsilon_2 > 0$. So, all 2×2 hypothesis matrices $(0 \le \lambda_i^j \le 1)$ may be divided into the following disjoint classes:

$$\begin{aligned} &(\mathrm{O})\ \ \varepsilon_1 \cdot \varepsilon_2 \ge 0, \\ &(\mathrm{A})\ \ \varepsilon_1 < 0,\ \varepsilon_2 > 0,\ \delta^1 < \delta^2 < 0, \\ &(\mathrm{B})\ \ \varepsilon_1 < 0,\ \varepsilon_2 > 0,\ \delta^1 < \delta^2 = 0, \qquad (3.58) \\ &(\mathrm{C})\ \ \varepsilon_1 < 0,\ \varepsilon_2 > 0,\ 0 < \delta^2 < -\delta^1, \\ &(\mathrm{D})\ \ \varepsilon_1 < 0,\ \varepsilon_2 > 0,\ 0 < \delta^2 = -\delta^1. \end{aligned}$$

Denote by $\pi_{(\nu)} = \{\pi_s(\xi),\ s = 1,2,\dots,\nu\}$ a uniformly optimal Markov strategy for the problem of loss minimization on the interval $(0,\nu)$, where $\nu < \infty$, and s is the *time remaining.*

In Case O there is a column all of whose elements are not less than the elements of the other column and there obviously exists an optimal strategy prescribing the constant use of the corresponding control (device). The description of optimal strategies for the remaining Cases A–D is similar, the difference between them consisting in the behaviour of the *a posteriori* probability.

Theorem 3.2 *There exists a sequence l_n, $n = 1,2,\dots$, such that $0 < l_n < 1$ and for any $\nu < \infty$ any optimal strategy for the problem of loss minimization over the time interval to ν for Cases A–D has the following form:*

$$\pi_s(\xi) := \begin{cases} 1 & \text{for } 0 \le \xi < l_s \\ 0 & \text{for } l_s < \xi \le 1 \\ \text{arbitrary} & \text{for } \xi = l_s, \end{cases}$$

where $s = 1,2,\dots,\nu$.

Proof. According to Theorem 2.4 it is sufficient to show the existence of sequences $\{l_n\}$ such that for all $s = 0, 1, \ldots$

$$TW_s(\xi) := \min\{T^1W_s(\xi), T^2W_s(\xi)\} = \begin{cases} T^1W_s(\xi) & \text{if } 0 \le \xi \le l_{s+1} \\ T^2W_s(\xi) & \text{if } l_{s+1} \le \xi \le 1. \end{cases} \tag{3.59}$$

To show this, it is sufficient in turn to check that the functions

$$r_{s+1}(\xi) := T^2W_s(\xi) - T^1W_s(\xi) \tag{3.60}$$

are continuous, strictly decreasing with respect to ξ and $r_{s+1}(0) > 0$, $r_{s+1}(1) < 0$ for $s = 0, 1, \ldots$. Then l_s may be taken as the zero of the function $r_s(\xi)$.

Continuity of $r_s(\xi)$ follows from the continuity of $W_\nu(\xi)$ for $\nu < \infty$ (proved at the end of §2.3) and the form of the operators T^1 and T^2 (see (3.57) and (3.55)).

Further, it is convenient to conduct the proof in the coordinates $\eta := \tilde{\eta}(\xi) := \ln[c\xi/(1-\xi)]$, where $c := -\varepsilon_1/\varepsilon_2$. Actually, the value of c in this transformation may be taken arbitrarily. As will be shown below, however, the value $c = -\varepsilon_1/\varepsilon_2$ is taken to obtain the equality $\tilde{\eta}(l_1) = 0$.

For reasons of simplicity, *assume* that all λ_i^j are different from 0 and 1. Then, if $\xi(0)$ differs from 0 and 1, $\xi(n)$ differs from 0 and 1 for all n and, correspondingly, $\tilde{\eta}(\xi)$ takes only finite values. Cases in which λ_i^j may be equal to 0 or 1 are considered similarly, but in such cases $\tilde{\eta}(\xi)$ may take the values $-\infty, +\infty$. The inverse transformation to $\tilde{\eta}(\xi)$ is denoted by

$$\tilde{\xi}(\eta) := e^\eta/(e^\eta + c). \tag{3.61}$$

Further, let

$$\begin{aligned} \widetilde{W}_s(\eta) &:= W_s(\tilde{\xi}(\eta)), & \tilde{r}_s(\eta) &:= r_s(\tilde{\xi}(\eta)) \\ \tilde{p}^j(\eta) &:= p^j(\tilde{\xi}(\eta)), & \tilde{q}^j(\eta) &:= q^j(\tilde{\xi}(\eta)), \end{aligned} \tag{3.62}$$

and designate by $\widetilde{T}^j$, $\widetilde{T}$ and $\widetilde{M}^j$ respectively the operators acting on functions depending on η corresponding to T^j, T and M^j. Then

$$\widetilde{M}^j f(\eta) := \tilde{p}^j(\eta) f(\eta + \gamma^{1j}) + (1 - \tilde{p}^j(\eta)) f(\eta + \gamma^{0j}), \tag{3.63}$$

where

$$\gamma^{1j} := \ln(\lambda_1^j/\lambda_2^j), \quad \gamma^{0j} := \ln\left[(1-\lambda_1^j)/(1-\lambda_2^j)\right] \tag{3.64}$$

and

$$\widetilde{T}^j f(\eta) := \tilde{q}^j(\eta) + \widetilde{M}^j f(\eta). \tag{3.65}$$

First we formulate a lemma about the properties of the operators $\widetilde{T}^j$, T^j, $\widetilde{M}^j$, M^j.

Lemma 3.1 *The following properties hold:*

(1) $M^j(b\xi + d) = b\xi + d$.

(2) If $f(\xi)$ is convex, then $M^j f(\xi) \le f(\xi)$.

(3) $\widetilde{M}^j(f - g) = \widetilde{T}^j f - \widetilde{T}^j g$ *for any f and g.*

(4) $\widetilde{M}^1\widetilde{M}^2 f = \widetilde{M}^2\widetilde{M}^1 f, \quad \widetilde{T}^1\widetilde{T}^2 f = \widetilde{T}^2\widetilde{T}^1 f$.

(5) If $f(\eta)$ is a nonincreasing function, then $\widetilde{M}^1 f(\eta)$ is also a nonincreasing function. Moreover, if $f(\eta)$ is continuous and $f(\eta + \gamma^{11}) \ne f(\eta + \gamma^{01})$, then $\widetilde{M}^1 f(n)$ is strictly decreasing in some neighbourhood of η. Similar statements replacing the words "nonincreasing, decreasing" with "nondecreasing, increasing" and with the condition $f(\eta+\gamma^{12}) \ne f(\eta+\gamma^{02})$ can be applied to the operator $\widetilde{M}^2$.

Properties 3 and 4 and a property similar to 5 hold also for the operators T^j and M^j.

Proof. Properties 1–4 may be checked directly from the definition of the operators M^j and T^j (see formulae (3.55)–(3.57)). Properties 1 and 2 reflect the fact that the process $\xi(n)$ is a martingale and Property 4 has a clear interpretation: the result of the sequential use of the first and the second devices does not depend on the order of their use. This is proved directly by virtue of the easily checked equality $\tilde{p}^1(\eta)\tilde{p}^2(\eta + \gamma^{11}) = \tilde{p}^2(\eta)\tilde{p}^1(\eta + \gamma^{12})$. (This last relation is a special case of the elementary formula $P\{A \cap B\} = P\{A\}P\{B|A\} = P\{B\}P\{A|B\}$.)

To prove Property 5, we write the following equation (for $\eta < \eta'$):

$$\begin{aligned}
&\widetilde{M}^j f(\eta') - \widetilde{M}^j f(\eta) \\
&\quad = \tilde{p}^j(\eta')f(\eta' + \gamma^{1j}) + (1 - \tilde{p}^j(\eta'))f(\eta' + \gamma^{0j}) \\
&\qquad - \tilde{p}^j(\eta)f(\eta + \gamma^{1j}) - (1 - \tilde{p}^j(\eta))f(\eta + \gamma^{1j}) \\
&\quad = (\tilde{p}^j(\eta') - \tilde{p}^j(\eta))((f(\eta' + \gamma^{1j}) - f(\eta' + \gamma^{0j})) \\
&\qquad + (\tilde{p}^j(\eta)(f(\eta' + \gamma^{1j}) - f(\eta + \gamma^{1j})) \\
&\qquad + (1 - \tilde{p}^j(\eta))(f(\eta' + \gamma^{0j}) - f(\eta + \gamma^{0j})).
\end{aligned} \tag{3.66}$$

Let $j = 1$. In Cases A–D, it follows from $\delta^1 < 0$ that $\gamma^{11} < 0$, $\gamma^{01} > 0$. Moreover, from $\delta^1 < 0$, formula (3.55) for $p^1(\xi)$ and the monotonicity of the transformation (3.61), it follows that $\tilde{p}^1(\eta)$ is strictly decreasing. Therefore if $f(\eta)$ is a nonincreasing function then each of the three terms in the last sum of (3.66) is nonpositive. If, moreover, $f(\eta)$ is continuous and $f(\eta + \gamma^{11}) \neq f(\eta + \gamma^{01})$, then this inequality also holds for η' which are near enough to η, and this means that the second term on the right-hand side of (3.66) is strictly positive. This proves Statement 5 for $j = 1$.

If $j = 2$, then in Cases C and D we obtain similarly that $\gamma^{12} > 0$, $\gamma^{02} < 0$ and $\tilde{p}^2(\eta)$ is strictly increasing and in Case A we obtain that $\gamma^{12} < 0$, $\gamma^{02} > 0$ and $\tilde{p}^2(\eta)$ is strictly decreasing. In Case B, $\widetilde{M}^2 f(\eta) \equiv f(\eta)$ holds. In all cases, if $f(\eta)$ is a nondecreasing function all terms in the last sum of (3.66) are nonnegative. If $f(\eta)$ is continuous and $f(\eta+\gamma^{12}) \neq f(\eta+\gamma^{02})$ (which is only possible for $\delta^2 \neq 0$), then all derivations are similar to the case with $j = 1$. The lemma is proven. ■

Now we return to the proof of Theorem 3.2. As was mentioned, it is sufficient to check that the function $\tilde{r}_s(\gamma)$ is strictly decreasing with respect to η and changes its sign. The proof of this fact is by induction.

For $s = 1$ we have, by the equality $\widetilde{W}_0(\eta) :\equiv 0$ and the relations (3.56), (3.55) and (3.53), that

$$\begin{aligned}
\tilde{r}_1(\eta) &= \tilde{T}^2\widetilde{W}_0(\eta) - \tilde{T}^1\widetilde{W}_1(\eta) \\
&= \tilde{q}^2(\eta) - \tilde{q}^1(\eta) = \tilde{p}^2(\eta) - \tilde{p}^1(\eta) \\
&= \tilde{\xi}(\eta)(\varepsilon_1 - \varepsilon_2) + \varepsilon_2.
\end{aligned}$$

Therefore, $\tilde{r}_1(\eta)$ is strictly decreasing with respect to η and changes sign, since $\tilde{\xi}(\eta)$ is monotonic (see (3.61)), and $\varepsilon_1 < 0$, $\varepsilon_2 > 0$. Moreover, $\tilde{r}_1(0) = 0$, so $l_1 = \tilde{\xi}(0) = -\varepsilon_2/(\varepsilon_1 - \varepsilon_2)$.

By Statement 4 of Lemma 3.1 and the equality $\widetilde{W}_n(\eta) = \tilde{T}\widetilde{W}_{n-1}(\eta)$ we have that

$$\begin{aligned}
\tilde{r}_{n+1}(\eta) &= \tilde{T}^2\widetilde{W}_n(\eta) - \tilde{T}^1\widetilde{W}_n(\eta) \\
&= (\tilde{T}^2\tilde{T}\widetilde{W}_{n-1}(\eta) - \tilde{T}^2\tilde{T}^1\widetilde{W}_{n-1}(\eta)) \\
&\quad + (\tilde{T}^1\tilde{T}^2\widetilde{W}_{n-1}(\eta) - \tilde{T}^1\tilde{T}\widetilde{W}_{n-1}(\eta)).
\end{aligned}$$

Using Statement 3 of Lemma 3.1, this expression for $\tilde{r}_{n+1}(\eta)$ may be written in the form

$$\begin{aligned}
&\tilde{r}_{n+1}(\eta) \\
&\quad = \widetilde{M}^2(\tilde{T}\widetilde{W}_{n-1}(\eta) - \tilde{T}^1\widetilde{W}_{n-1}(\eta)) - \widetilde{M}^1(\tilde{T}\widetilde{W}_{n-1}(\eta) - \tilde{T}^2\widetilde{W}_{n-1}(\eta)).
\end{aligned}$$

From this, using the optimality equation (3.59), we obtain

$$\tilde{r}_{n+1}(\eta) = \widetilde{M}^1[\tilde{r}_n(\eta)]^+ - \widetilde{M}^2[\tilde{r}_n(\eta)]^-, \tag{3.67}$$

where $a^+ := \max(a, 0)$, $a^- := \max(-a, 0)$.

Suppose the result is proved for $s = n$, i.e. that $\tilde{r}_n(\eta)$ is strictly decreasing with respect to η and changes sign at the point $\eta_n := \tilde{\eta}(l_n)$. Then from Statement 5 of Lemma 3.1 and from (3.63) it follows that $\widetilde{M}^1[r_n(\eta)]^+$ is strictly decreasing for $\eta < \eta_n - \gamma^{11}$ and equal to 0 for $\eta \geq \eta_n - \gamma^{11}$. Similarly, $\widetilde{M}^2[\tilde{r}_n(\eta)]^-$ is nondecreasing for all η, strictly increasing for $\eta > \eta_n - \max\{\gamma^{02}, \gamma^{12}\}$ and equal to 0 for $\eta \leq \eta_n - \max\{\gamma^{02}, \gamma^{12}\}$. From (3.58) it is not difficult to see that in all cases $\gamma^{11} < \gamma^{12}$. Therefore from (3.67) and what was said above it follows that $\tilde{r}_{n+1}(\eta)$ is strictly decreasing and for some $\eta := \eta_{n+1}$ changes sign.

Theorem 3.2 is proven. ■

Remark 3.2 If $\delta_2 = 0$, i.e. we have the Bellman case, then $\gamma^{02} = \gamma^{12} = 0$ and it follows directly from the proof that the sequence l_n increases. ■

Remark 3.3 If the matrix $\{\lambda_i^j\}$, $i, j = 1, 2$, is symmetric, i.e. $\lambda_1^1 = \lambda_2^2$, $\lambda_2^1 = \lambda_1^2$, then it is easy to check that $l_n \equiv 1/2$ and the optimal strategy becomes stationary, i.e. it does not depend on the number of trials remaining. ■

Remark 3.4 Suppose the cost function $q_k(\xi)$ depends on the number of tests and specifically is given by the (original) cost (3.56) multiplied by c_k, where $\{c_k\}$ is a nonincreasing sequence. It is easy to check that in this case the statement of the theorem remains true. (The proof is similar to the above with appropriate changes to Property 4 of Lemma 3.1.) ■

* * *

Consider now the problem on an infinite time interval. If Λ is an F-matrix (i.e. either $\lambda_1^2 \neq \lambda_2^2$, or $\lambda_1^2 = \lambda_2^2$ and $\lambda_2^1 = 1$), then as $s \to \infty$ we have $W_s(\xi) \to W(\xi) < \infty$ and we may take limits in the optimality equation (3.59). In this case the functions $r_s(\xi)$ given by (3.60) also converge pointwise to some limit $r(\xi)$. The limit function $r(\xi)$ is continuous for $0 < \xi < 1$ by the continuity $W(\xi)$ (see Theorem 2.3) and does not decrease with respect to ξ, as the limit of a sequence of nondecreasing functions. Let l_∞ be the limit point of the sequence $\{l_s\}$. By Theorems 2.5 and 2.3 the following theorem holds.

Theorem 3.3 *For F-matrices the strategy given by the functions*

$$\pi_s(\xi) := \begin{cases} 1 & \text{if } 0 \leq \xi < l_\infty \\ 0 & \text{if } l_\infty < \xi < 1 \\ \text{arbitrary} & \text{if } \xi = l_\infty \end{cases}$$

is optimal for the problem of loss minimization on an infinite time interval. ■

We note here that l_∞ is a point at which the limit function $r(\xi)$ changes sign. If we can prove that, similarly to all $r_s(\xi)$, $r(\xi)$ has a single sign change, then it will follow that the strategy given in

Theorem 3.3 is the *unique* optimal strategy. We will not prove this fact, but will only indicate one of the possible ways to prove it. It is possible to calculate the value of loss corresponding to the *stationary threshold* strategy, i.e. the strategy given by the function

$$\pi^l(\xi) := \begin{cases} 1 & \text{if } 0 \leq \xi < l \\ 0 & \text{if } l < \xi \leq 1 \\ \text{arbitrary} & \text{if } \xi = l \end{cases}$$

(a calculation method is given, for example, in Reimnitz 1977). Then, amongst thresholds we need to choose one that gives minimal loss and prove that corresponding to this threshold the value function and the corresponding threshold strategy satisfy the optimality equation and that the function $r(\xi)$ changes the sign at this point. We shall use such an approach to prove the corresponding theorem in continuous time.

3.5 Scheme with sharable resources

We give a formal description more general than that of the basic scheme control problem of §2.1, in the framework of which we will formulate the scheme with *sharable resources* discussed in Chapter 1. The difference here from §2.1 is that now the control space is *not* a set of *vertices* of the simplex, but the *simplex* itself (for $N = 2$ it is an interval) and the observation space is expanded by a vector b showing the state of the devices (i.e. which devices are switched on and which are switched off). For simplicity, we consider only the case $m = N = 2$.

On a measurable space $(\Omega, \mathcal{F})$ suppose given the following:

(a) A random *parameter value* θ with values in $\widehat{S}^2$, corresponding to the two hypotheses.

(b) A one-dimensional random process $a := \{a(n,\omega),\ n = 1, 2, \ldots\}$, where $a(n) := a(n,\omega)$ takes values in $[0,1]$ and is interpreted in the following way. Two *independent* trials are conducted at time n. The probability of a positive result for the first trial equals $a^1(n) := a(n)$ and that for the second trial equals $a^2(n) := 1 - a(n)$. If the result of the first (correspondingly, the second) trial is positive (1), then an observation is taken on the first (second) device at time n (the device is switched *on*), and in the case of a negative result (0) the observation is not taken (the device is swtiched *off*). Note that an observation may be taken on *either* one of the devices, or on *both* or *no* observation at all.

(c) A two-dimensional random process $b := \{b(n,\omega),\ n = 1,2,\dots\}$ and another $X' := \{X'(n,\omega),\ n = 1,2,\dots\}$. Here $b(n) := b(n,\omega) := (b^1(n), b^2(n))$ takes one of four possible values $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ and corresponds to the results of the two trials described above conducted at time n, i.e. $b^j(n) = 1$ corresponds to the fact that the j^{th} device is switched on at time n. Thus if $a^j(n) = 0$, then $b^j(n) = 0$, $j = 1,2$. For the process X', we let $X'(0,\omega) := (0,0)$ and define $X'(n) := X'(n,\omega) := (X'^1(n), X'^2(n))$, where $X'^j(n)$ corresponds to the number of successes observed up to time n on the j^{th} device; here $\Delta X'(n) := (\Delta X'^1(n), \Delta X'^2(n)) := X'(n) - X'(n-1)$ takes values $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ and, if $b^j(n) = 0$, then $\Delta X'^j(n) = 0$, $j = 1,2$.

Analogously to §2.1, the following σ-algebras are defined for $n \geq 0$:

$$\begin{aligned}
\mathcal{F}'_n &:= \sigma(a(s), b(s), \Delta X'(s),\ 1 \leq s \leq n),\\
\widetilde{\mathcal{F}}'_{n+1} &:= \mathcal{F}'_n \vee \sigma(a(n+1)),\\
\widehat{\mathcal{F}}'_{n+1} &:= \widetilde{\mathcal{F}}'_{n+1} \vee \sigma(b(n+1)),\\
\mathcal{F}'^{\theta}_n &:= \mathcal{F}'_n \vee \sigma(\theta),\\
\widetilde{\mathcal{F}}'^{\theta}_{n+1} &:= \widetilde{\mathcal{F}}'_{n+1} \vee \sigma(\theta),\\
\widehat{\mathcal{F}}'^{\theta}_{n+1} &:= \widehat{\mathcal{F}}'_{n+1} \vee \sigma(\theta).
\end{aligned}$$

Now at each time $n = 1,2,\dots$ a *nonrandomized* action rule β', depending on the realizations of previous controls (to which the process $a(n,\omega)$ corresponds) and previous observations (to which the processes $b(n)$ and $\Delta X'(n)$ correspond), define the probabilities for conducting the next pair of trials. A (*randomized*) action rule is a sequence of random measures $\{\beta'_n(d\alpha,\omega),\ n = 1,2,\dots\}$ on the interval $[0,1]$, where $\beta'(d\alpha) := \beta'(d\alpha,\omega)$ is $\mathcal{F}'_{n-1}$-measurable for each $n \geq 1$. We denote by Π' the set of such action rules. Given $\xi \in [0,1]$, corresponding to the *a*

priori probability of the first hypothesis, and $\beta' \in \Pi'$, a measure $P^{\beta'}_{\xi}$ is defined on $\mathcal{F}'_{\infty}$ by the following formulae (compare with (2.2)–(2.4)):

$$\begin{aligned}
P^{\beta'}_{\xi}\{\theta_1 = 1\} &= \xi,\\
P^{\beta'}_{\xi}\{\theta_2 = 1\} &= 1 - \xi,\\
P^{\beta'}_{\xi}\{a(n) \in A | \mathcal{F}'^{\theta}_{n-1}\} &= \beta'_n(A), \qquad (3.68)\\
P^{\beta'}_{\xi}\{b^j(n) = 1 | \widetilde{\mathcal{F}}'^{\theta}_n\} &= a^j(n),\\
P^{\beta'}_{\xi}\{\Delta X'^j(n) = 1 | \widehat{\mathcal{F}}'^{\theta}_n\} &= \Big(\sum_{i=1}^{2} \lambda^j_i \theta_i\Big) b^j(n),
\end{aligned}$$

together with the property of *conditional independence* of the coordinates of the vector $b(n)$ (correspondingly $\Delta X(n)$) when conditioned on $\widetilde{\mathcal{F}}'^{\theta}_n$ (correspondingly $\widehat{\mathcal{F}}'^{\theta}_n$).

Loss up to time $\nu \leq \infty$ again has the form (compare with (2.8))

$$\begin{aligned}
W^{\beta'}_{\nu}(\xi) &= \nu(\lambda_1 \xi + \lambda_2(1-\xi)) - E^{\beta'}_{\xi}\left[X'^1(\nu) + X'^2(\nu)\right]\\
&= \sum_{n=1}^{\nu} E^{\beta'}_{\xi} \theta(\tilde{\Lambda} - \Lambda) a^*(n).
\end{aligned}$$

Define $W'_{\nu}(\xi) := \inf_{\beta' \in \Pi'} W^{\beta'}_{\nu}(\xi)$ for $\nu \leq \infty$.

Now we are in a position to describe the scheme with *sharable resources*. It results from that described above by excluding from the observation space the vector $b(n,\omega)$ (i.e. by restricting the σ-algebras $\mathcal{F}'_n$ and $\widetilde{\mathcal{F}}'_n$) and make the corresponding changes in the definition of the action rules.

Let

$$\begin{aligned}
\mathcal{F}''_n &:= \sigma(a(s), \Delta X'(s),\ 1 \leq s \leq n),\\
\widetilde{\mathcal{F}}''_{n+1} &:= \mathcal{F}''_n \vee \sigma(a(n+1)).
\end{aligned}$$

Consideration of this σ-algebra means that (unlike the case for the σ-algebras $\mathcal{F}'_n$ and $\widetilde{\mathcal{F}}'_n$) we do not discriminate between an absence of success at some time connected with the fact that the device was switched off and a similar event connected with the fact that the result of the observation was a failure.

Consider the *action rules* $\beta'' := \{\beta''_n(d\alpha), n = 1,2,\ldots\}$, where $\beta''_n(d\alpha)$ is an $\mathcal{F}''_{n-1}$-measurable random measure on the interval $[0,1]$, $n = 1,2,\ldots$, and denote the set of such action rules by Π''. The measure $P_\xi^{\beta''}$ on $\mathcal{F}''_\infty$ is defined by formulae similar to (3.68), except that in the third formula $\mathcal{F}'^{\theta}_{n-1}$ is replaced by $\mathcal{F}''^{\theta}_{n-1}$, and the last two formulae are replaced by

$$P_\xi^{\beta''}\{\Delta X'^j(n) = 1 | \widetilde{\mathcal{F}}''^{\theta}_n\} = (\sum_{i=1}^N \lambda_i^j \theta_i) a^j(n). \qquad (3.69)$$

In this case $a^j(n,\omega)$, $j = 1,2$ may be considered as the *proportion* of a unit *resource* allocated to the j^{th} device, since according to (3.69) under the i^{th} hypothesis as a result of the use of a control α in $[0,1]$ the probability of a success on the j^{th} device equals $\alpha^j\lambda_i^j$, where $\alpha^1 := \alpha$, $\alpha^2 := 1-\alpha$.

It is obvious that any action rule $\beta'' \in \Pi''$ belongs also to Π', and the corresponding measures coincide on $\mathcal{F}''^{\theta}_\infty$. Therefore, if for $\nu \le \infty$ (and $\beta'' \in \Pi''$) we define

$$W_\nu^{\beta''}(\xi) = \sum_{n=1}^{\nu} E_\xi^{\beta''}\left[\theta(\tilde\Lambda - \Lambda)a^*(n)\right],$$
$$W''_\nu(\xi) = \inf_{\beta''\in\Pi''} W_\nu^{\beta''}(\xi),$$

then, obviously, $W''_\nu(\xi) \ge W'_\nu(\xi)$. Moreover, each action rule β for the basic scheme may be considered as an action rule in the scheme with sharable resources (i.e. from Π''), such that the corresponding measures $\beta''(d\alpha)$, $n = 1,2,\ldots$, are concentrated on two points $\alpha = 0$ and $\alpha = 1$ and the loss values of the two models coincide. At the same time the set of *action rules* for the scheme with sharable resources is significantly larger than that for the basic scheme. So,

$$W_\nu(\xi) \ge W''_\nu(\xi) \ge W'_\nu(\xi). \qquad (3.70)$$

We show that the possibility of sharing resources does not cause the loss functional to decrease.

Theorem 3.4 *For any* ν $(1 \le \nu \le \infty)$

$$W_\nu(\xi) = W'_\nu(\xi). \qquad (3.71)$$

Proof. As in Chapter 2, it is convenient for the proof of the theorem to transfer study to the process $\xi(n,\omega)$ corresponding to the *a posteriori* probability of the first hypothesis. As is shown in Chapter 2, a consideration of the problem in terms of this process is equivalent to the study of the following homogeneous Markov model.

The points ξ of the interval $[0,1]$ are the states and the controls are also points α in $[0,1]$. Using the control $\alpha \in [0,1]$ the point ξ makes a transition with probability $(1-\alpha)\alpha$ to the same point ξ. This corresponds to the situation in the initial problem in which both devices are switched off. With probability $\alpha\cdot\alpha\cdot p^1(\xi)$ (correspondingly $\alpha\cdot\alpha\cdot(1-p^1(\xi))$) a transition to the point $\Gamma^{11}\xi$ (correspondingly $\Gamma^{01}\xi$) occurs. This corresponds to the observation of a success (failure) on the first device when the second device is switched off. Analogously, the transition to the points $\Gamma^{12}\xi$ and $\Gamma^{02}\xi$ occurs with probabilities $(1-\alpha)(1-\alpha)p^2(\xi)$ and $(1-\alpha)(1-\alpha)(1-p^2(\xi))$ respectively. (Here the function $p^j(\xi)$ and the transformation $\Gamma^{ij}\xi$ $(j = 1,2;\ i = 0,1)$ is defined by formula (3.55).) The situation in which both devices are switched on and two successes are observed corresponds to a transition to the point $\Gamma^{11}\Gamma^{12}\xi$, which occurs with probability

$$\alpha(1-\alpha)[\lambda_1^1\lambda_1^2\xi + \lambda_2^1\lambda_2^2(1-\xi)] = \alpha(1-\alpha)p^1(\xi)p^2(\Gamma^{11}\xi)$$
$$= \alpha(1-\alpha)p^2(\xi)p^1(\Gamma^{12}\xi). \qquad (3.72)$$

The equalities in (3.72) follow directly from (3.55). (See also the proof of Property 4 of Lemma 3.1.) Similarly, the transition to the points $\Gamma^{11}\Gamma^{02}\xi$, $\Gamma^{01}\Gamma^{02}\xi$ and $\Gamma^{01}\Gamma^{12}\xi$ occur with the respective probabilities

$$\alpha(1-\alpha)p^1(\xi)(1-p^2(\Gamma^{11}\xi)),$$
$$\alpha(1-\alpha)(1-p^1(\xi))(1-p^2(\Gamma^{01}\xi)),$$
$$\alpha(1-\alpha)(1-p^1(\xi))(p^2(\Gamma^{01}\xi)).$$

The cost function in the corresponding model is given by

$$q^\alpha(\xi) := \alpha q^1(\xi) + (1-\alpha)q^2(\xi),$$

where $q^j(\xi)$, $j = 1,2$, is defined in (3.56). It is not difficult to check that the action of the operator T^α corresponding to the control

$\alpha \in [0,1]$ is given by the formula

$$T^\alpha f(\xi) = \alpha \cdot \alpha T^1 f(\xi) + (1-\alpha)(1-\alpha)T^2 f(\xi)$$
$$+ \alpha(1-\alpha)[q^1(\xi) + q^2(\xi) + f(\xi) + M^1 M^2 f(\xi)], \quad (3.73)$$

where the operators T^j and M^j, $j = 1, 2$, are defined in (3.57).

Since we have constructed a model (in initial form) with discrete transitions (see §2.4), by Theorem 2.3 the optimality equation holds, i.e.

$$W'_{s+1}(\xi) = \min_{0\le\alpha\le 1} T^\alpha W'_s(\xi), \quad (3.74)$$

where s denotes time remaining.

Note that the optimality equation for $W_{s+1}(\xi)$ may be rewritten as

$$W_{s+1}(\xi) = \min_{0\le\alpha\le 1}\left[\alpha T^1 W_s(\xi) + (1-\alpha)T^2 W_s(\xi)\right]$$
$$:= T_s W_s(\xi) := q_s(\xi) + M_s W_s(\xi), \quad (3.75)$$

where $q_s(\xi) := q^1(\xi)$ if $0 \le \xi \le l_{s+1}$, $q_s(\xi) := q^2(\xi)$ if $l_{s+1} \le \xi \le 1$,

$$M_s f(\xi) := \begin{cases} M^1 f(\xi) & \text{if } 0 \le \xi \le l_{s+1} \\ M^2 f(\xi) & \text{if } l_{s+1} < \xi \le 1, \end{cases} \quad (3.76)$$

and the sequence $\{l_s\}$ is defined in Theorem 3.2.

We will also need the following properties of the operator M_s, $s = 1, 2$, which are true from (3.76) because they hold for M^1 and M^2 (see Lemma 3.1):

$$M_s[f(\xi) + g(\xi)] = M_s f(\xi) + M_s g(\xi), \quad (3.77)$$
$$M_s[f(\xi) - g(\xi)] = T_s f(\xi) - T_s g(\xi), \quad (3.78)$$
$$M_s[c\xi + d] = c\xi + d. \quad (3.79)$$

Now we return to the proof of (3.71) for $\nu < \infty$. It is obvious that

$$W'_1(\xi) = \min\left[q^1(\xi), q^2(\xi)\right] = W_1(\xi), \quad (3.80)$$

so (3.71) holds for $\nu = 1$. Let (3.71) be proven for all $\nu \le s$. Then for $\nu = s+1$, taking account of (3.73), relation (3.74) may be rewritten as

$$W'_{s+1}(\xi) = W_{s+1}(\xi) + \min_{0\le\alpha\le 1}\Big\{\left[\alpha \cdot \alpha(T^1 W_s(\xi) - W_{s+1}(\xi))\right]$$
$$+ (1-\alpha)(1-\alpha)\left[T^2 W_s(\xi) - W_{s+1}(\xi)\right]$$
$$+ \alpha(1-\alpha)\left[q^1(\xi) + q^2(\xi) + M^1 M^2 W_s(\xi)\right.$$
$$\left. + W_s(\xi) - 2W_{s+1}(\xi)\right]\Big\}. \quad (3.81)$$

By (3.75) the first and second terms within the minimization of (3.81) are nonnegative. If it is possible to prove that the third term is also nonnegative, then taking account of (3.70), it follows from (3.81) that $W'_{s+1}(\xi) = W_{s+1}(\xi)$.

So, to prove (3.71) it is sufficient to show that for all $s = 1, 2, \ldots$

$$A_s(\xi) = q^1(\xi) + q^2(\xi) + M^1 M^2 W_s(\xi) + W_s(\xi) - 2W_{s+1}(\xi) \ge 0. \quad (3.82)$$

This relation will also be proved by induction. Since $W_0(\xi) \equiv 0$, then from (3.80) it follows that $A_0(\xi) \ge 0$. For $t \le s$, suppose it is proven that $A_t(\xi) \ge 0$.

Applying the operator M_s to (3.82), using (3.77) and (3.79) sequentially and subtracting the result from $A_{s+1}(\xi)$, we have that

$$A_{s+1}(\xi) - M_s A_s(\xi)$$
$$= q^1(\xi) + q^2(\xi) + M^1 M^2 W_{s+1}(\xi) + W_{s+1}(\xi)$$
$$- 2W_{s+2}(\xi) - q^1(\xi) - q^2(\xi) - M_s M^1 M^2 W_s(\xi)$$
$$- T_s W_s(\xi) + 2T_s W_{s+1}(\xi). \quad (3.83)$$

Now using the equality $W_{s+1}(\xi) = T_s W_s(\xi)$ (see (3.75)), from (3.83) we have that

$$A_{s+1}(\xi) = M_s A_s(\xi) + 2\left[T_s W_{s+1}(\xi) - W_{s+2}(\xi)\right]$$
$$+ \left[M^1 M^2 W_{s+1}(\xi) - M_s M^1 M^2 W_s(\xi)\right]. \quad (3.84)$$

The first term of the right-hand side of (3.84) is nonnegative by the induction assumption and the fact that $f \geq 0$ implies $M_s f \geq 0$. The second term is nonnegative, since (see (3.75) and (3.76))

$$T_s W_{s+1}(\xi) \geq T W_{s+1}(\xi) = \min\left[T^1 W_{s+1}(\xi), T^2 W_{s+1}(\xi)\right] = W_{s+2}(\xi),$$

which means that to prove (3.82) it is sufficient to check that $M^1 M^2 W_{s+1}(\xi) - M_s M^1 M^2 W_s(\xi) \geq 0$ for any $s = 1, 2, \ldots$. But to show this, it is sufficient to show that for $j = 1, 2$,

$$M^1 M^2 W_{s+1}(\xi) \geq M^j M^1 M^2 W_s(\xi). \tag{3.85}$$

By Lemma 3.1, the operators M^1 and M^2 are commutative, therefore to prove (3.85) it is sufficient to show that

$$W_{s+1}(\xi) \geq M^j W_s(\xi), \qquad j = 1, 2, \quad s = 1, 2, \ldots. \tag{3.86}$$

By Theorem 2.2, the function $W_s(\xi)$ is convex, which means that according to Property 2 of Lemma 3.1, $W_s(\xi) \geq M^j W_s(\xi)$. But for all s, from the nonnegativity of the cost functions, $W_{s+1}(\xi) \geq W_s(\xi)$. Therefore (3.86) is proven and hence (3.71) also for $\nu < \infty$.

From this, $\lim_{\nu \to \infty} W_\nu(\xi) = \lim_{\nu \to \infty} W'_\nu(\xi)$. But according to Theorem 3.1, for problems of the basic scheme $\lim W_\nu(\xi) = W_\infty(\xi)$. From the inequality $W'_\infty(\xi) \geq \lim_{\nu \to \infty} W'_\nu(\xi)$ (see (2.49) and (3.70)) it follows that (3.71) holds also for $\nu = \infty$. ■

4 PROBLEM FORMULATION AND SOLUTION METHODS IN CONTINUOUS TIME

4.1 Reduction of continuous to discrete time

Chapter 2 gives a mathematical formulation of a situation which may be presented informally as follows. We have m devices, each of them generating independent Bernoulli (0 or 1) random variables with parameters depending on the number of devices. We have N hypotheses about parameter values and a given *a priori* distribution $\xi = (\xi_1, \ldots, \xi_N) \in S^N$ on the set of these hypotheses. At each moment of time only one of these devices may be used (i.e. the control action takes values in $\widehat{S}^m$) and the corresponding random variable may be observed. Accordingly, since the observations determine a payoff, it is required at each moment of time to decide which device should be used based on the data of which device was used at previous moments of time and the observations that were observed on each.

In the case of continuous time, the analogue of the *Bernoulli* process is the *Poisson* process. Therefore in continuous time it is natural to assume that under the i^{th} hypothesis the device with index j generates a Poisson process with parameter λ_i^j.

However, in continuous time it is not sufficient to take as controls functions taking values in $\widehat{S}^m$. A heuristic explanation of this fact was given in §1.6. Therefore, in continuous time the control $a(t)$ takes values in S^m, so that $a^j(t)$ corresponds to that *fraction* of a unit resource allocated to the j^{th} device at time t, and the *jump intensity* on this device under the i^{th} hypothesis equals $\lambda_i^j a^j(t)$.

In §4.2 a precise formulation of the continuous time problem based on the concepts of *martingale* and *point process* is given. The Appendix contains the facts necessary to understand this part. Moreover, in §4.2 it is shown that similarly to discrete time the problem may