

Article

# On Energy–Information Balance in Automatic Control Systems Revisited

Vladimir Rubtsov <sup>1,2,3</sup> 

<sup>1</sup> Département de Mathématiques, LAREMA UMR 6093 du CNRS, Université d'Angers, 49045 Angers, France; volodya@univ-angers.fr

<sup>2</sup> Theory Division, Institute for Theoretical and Experimental Physics, 117218 Moscow, Russia

<sup>3</sup> Institute of Geometry And Physics, 34151 Trieste, Italy

Received: 9 November 2020; Accepted: 13 November 2020; Published: 15 November 2020



**Abstract:** We revise and slightly generalize some variational problems related to the “informational approach” in the classical optimization problem for automatic control systems which was popular from 1970–1990. We find extremals for various degenerated (derivative independent) functionals and propose some interpretations of obtained minimax relations. The main example of such functionals is given by the Gelfand–Pinsker–Yaglom formula for the information quantity contained in one random process in another one. We find some balance relations in a linear stationary one-dimensional system with Gaussian signal and interpret them in terms of Legendre duality.

**Keywords:** Pinsker information rate; entropy rate; Legendre duality

*To Nikolay Petrovich Bukanov with my gratitude and admiration*

## 1. Introduction

An informational approach to an optimality criteria for a control system synthesis was proposed by Bukanov in [1]. He has considered the following one-dimensional, control system.

**Remark 1.** The “subtraction” and “addition” notation of the diagram nodes means that f.e. the input function  $X(t) = U(t) - P(t)$  and  $Z(t) = Y(t) - Y_0(t)$  while  $V(t) = Z(t) + N(t)$ .

His choice of an informational criterion was motivated by a straightforward engineering application: studies of the automatic board control system of the passenger jet TU-154. Mathematically, he had used the classical Gelfand–Pinsker–Yaglom functional formula [2]

$$J_{U,Z} = -\frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(1 - \mathcal{R}_{UZ}^2) d\omega$$

for the amount of information about the one random process (the defect or the error signal  $Z$ ) contained in another such process (the control signal  $U$ ), where  $\mathcal{R}_{UZ}$  denotes the mutual correlation function for these processes.

Pinsker had re-written this functional in the following form:

$$J_{U,Z} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{USZ}}{s_{USZ} - |s_{UZ}|^2} d\omega. \quad (1)$$

Here,  $S_U, S_Z$  are spectral densities of the powers of the control signal  $U$  and the defect signal  $Z$ ,  $S_{UZ}$  - is the mutual spectral density.

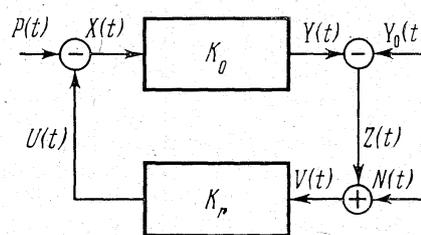
Bukanov has expressed the functional variables of (1) in terms of the object-regulator transfer functions and spectral densities of the perturbations and has obtained the following explicit integral:

$$J_{U,Z} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{(\mathcal{R}_0 s_P)^2 + (\mathcal{R}_r s_N)^2 + (1 + \mathcal{R}_0^2 \mathcal{R}_r^2) s_N s_P}{(1 + \mathcal{R}_0^2 \mathcal{R}_r^2 + 2\mathcal{R}_0 \mathcal{R}_r \cos(\phi_0 + \phi_r)) s_N s_P} d\omega. \tag{2}$$

The transfer functions of the object and the regulator are represented by the frequency responses  $K_0(j\omega) = \mathcal{R}_0 \exp(j\omega)$  and  $K_r(j\omega) = \mathcal{R}_r \exp(j\omega)$ . (We shall describe below the detailed reduction of the Pinsker Formula (1) to the Bukanov’s expression (2) in a slightly general context.)

The aim of this paper to revise and to generalize this optimal regulator problem (for a class of linear stationary systems with non-zero program ( $Y_0$ )) as the classical optimization problem for automatic control system of Figure 1. We get the following results:

- find extremals for a generalization of (1) and for some other degenerated (derivative independent) “informational” functionals (the main example of such functionals is given by the Pinsker formula for the entropy quantity of one random process in another.)
- obtained some minimax relations and propose its new interpretations as an “energy–balance equilibrium” in a spirit of the Brillouin’s and Schrodinger’s ideas of “negentropy” (we remind briefly this notion in the Discussion section).
- write our balance relations in a linear stationary one-dimensional system with Gaussian signals and interpret them in terms of Legendre duality.



**Figure 1.** This is a figure from [1]. Here,  $K_0$  denotes the operator object which transform the input function  $X(t)$  to the output function  $Y(t)$  which should be compared with the given function  $Y_0$  (the program function). The “defect” signal  $Z$  together with a measurement “noise”  $N$  forms the input of a regulator given by an operator  $K_r$ . The resulting control signal  $U$  with an external perturbation  $P$  goes back to the input of the object operator  $K_0$ .

We would like to stress and warn a mathematical purist reader that this manuscript is not a “mathematical paper” in a formal sense of this notion. We (almost) never discuss and almost never precise any “existence and finitude conditions”. We would rather prefer formal manipulations whenever they exist. Therefore, we decided not to overload these small notes with probability and measure theory precise statements, technicalities, and terminology. We refer all interested readers to the book [2] for all rigorous conditions and other pure mathematical details and statements.

### 1.1. The Origins

The famous Shannon’s formula [3] provides a facility to measure the capacity of a communication channel in a presence not only of *white noise* but also in the case when the transmission is perturbed by any Gaussian noise with the power spectrum function  $\epsilon(\omega)$  proportional to the square module of a filter transfer function  $y(\omega)$ :

$$\epsilon(\omega) = \kappa |y(\omega)|^2.$$

We recall it because it is probably the first when implicitly the informational characteristics are stipulated by some “energy” constraint. Namely, let the power of the “transmitter” is bounded by some (positive) quantity  $A^2$  such the spectral function of the transmitted signal is  $s(\omega)$ :

$$\int_{\omega_0}^{\omega_1} s(\omega) d\omega = A^2. \quad (3)$$

The Shannon formula we have mentioned above for the total capacity of  $s(\omega)$  reads for the frequency bandwidth  $[\omega_0, \omega_1]$  as

$$C_{s,\epsilon} = \int_{\omega_0}^{\omega_1} \log\left(1 + \frac{s(\omega)}{\epsilon(\omega)}\right) d\omega.$$

Now, it is easy to give a mathematical answer to the natural question: what should we do to get a maximal transmission rate for the given constraint (3)? Shannon proposes to consider the variational problem with the Lagrange multiplier:

$$\max\left\{\int_{\omega_0}^{\omega_1} \left[\log\left(1 + \frac{s(\omega)}{\epsilon(\omega)}\right) + \lambda s(\omega)\right] d\omega\right\} \quad (4)$$

which gives the condition

$$\lambda + \frac{1}{s(\omega) + \epsilon(\omega)} = 0$$

and implies that  $s(\omega) + \epsilon(\omega)$  is a constant. Going further, we put  $s(\omega) = -\frac{1}{\lambda} - \epsilon(\omega)$  in (3) and obtain

$$\lambda = -\frac{\omega_1 - \omega_0}{A^2 + B^2},$$

where  $B^2 = \int_{\omega_0}^{\omega_1} \epsilon(\omega) d\omega$  and the extremal  $s_{max}(\omega) = \frac{A^2 + B^2}{\omega_1 - \omega_0} - \epsilon(\omega)$ , while

$$C_{s,\epsilon}^{max} = (\omega_1 - \omega_0) \log \frac{A^2 + B^2}{\omega_1 - \omega_0} - \int_{\omega_0}^{\omega_1} \epsilon(\omega) d\omega.$$

In this paper, we are going to apply these beautiful observations of the founding father of Theory of Information to some (degenerated) variational problems with power constraints arising in the optimal control of automatic systems.

### 1.2. Review of Some Previously Known Results

The informational criteria for a control system synthesis were very popular in applied automatic control and measure systems in 1970–1980. We should mention here relevant to our interests works of academician Petrov’s school researchers [4]. They were concentrated on the questions of correctness and regularization (in sense of Tikhonov) the control and statistical optimization problems. Their most important and most interesting (to our aims) input has concluded in a study of a connection between the regularization and a correctness of problems in a one side and an approach motivated by the above Shannon channel capacity formulae in the other. Mathematically, their main tool was based on the Wiener filter theory and related integral equations of the Kolmogorov–Wiener type (see, for example, [5]).

It seems that the system considered by N. Bukanov was ideologically similar to models elaborated by Petrov’s school researchers [5,6]. In the same time, one should stress that Bukanov’s choice of the informational criterion has some advantages. Its mathematical toolbox does not appeal to integral equations and was strongly motivated as we have mentioned by engineering applications to numerical computations related with an optimization of the jet TU-154 automatic board control system.

We shall compare both approaches—[1,5,6] for some basic examples of control linear systems.

We remind that, in [1], a solution was proposed for the variational problem in the class of linear stationary systems with zero program ( $Y_0 = 0$ ) and Gaussian perturbation signal functions. Namely, Formula (1) makes sense exactly in this class of random functions admissible in [2].

More precisely, the main result of [1] is that the (degenerated) variational problem for the functional (2) has the following extremals:

$$\mathcal{R}_r = \mathcal{R}_0 \frac{s_P}{s_N}, \quad \phi_r = -\phi_0, \quad (5)$$

which are minimized at the same time as the functional of the error signal dispersion:

$$D_{Z,Z} = \int_{-\infty}^{+\infty} \frac{\mathcal{R}_0^2 (s_P + \mathcal{R}_r^2 s_N)}{1 + \mathcal{R}_0^2 \mathcal{R}_r^2 + 2\mathcal{R}_0 \mathcal{R}_r \cos(\phi_0 + \phi_r)} d\omega \quad (6)$$

and the functional of the control signal power dispersion:

$$D_{U,U} = \int_{-\infty}^{+\infty} \frac{\mathcal{R}_r^2 (s_N + \mathcal{R}_0^2 s_P)}{1 + \mathcal{R}_0^2 \mathcal{R}_r^2 + 2\mathcal{R}_0 \mathcal{R}_r \cos(\phi_0 + \phi_r)} d\omega. \quad (7)$$

One can conclude that the extremal relations (5) between the object–regulator characteristics guarantee that the optimal regulator provides the minimal power to the object of control. Another interesting observation made in [1] is that, in spite of the functional (2) minimal value  $(J_{U,Z})_{min} = 0$ , it does not mean that the regulator does not use the information at all—it means only that there is a certain “informational balance”:

$$J_{U,Z} = [C_{0,P,N} - Q_0] + [C_{r,P,N} - Q_r] \quad (8)$$

between the channel capacity

$$C_{0,P,N} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(1 + \mathcal{R}_0^2 \frac{s_P}{s_N}) d\omega.$$

of the object channel in the presence of perturbations  $P$  and  $N$ , the channel capacity

$$C_{r,P,N} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(1 + \mathcal{R}_r^2 \frac{s_N}{s_P}) d\omega.$$

of the regulator channel in the presence of perturbations  $P$  and  $N$ , and the (half-)differences

$$Q_0 = \frac{1}{2}(Q'_0 - Q_0''), \quad Q_r = \frac{1}{2}(Q'_r - Q_r'')$$

of the informational performances  $Q_0, Q_r$  of the output of the object and the regulator channels with open ( $Q'_0, Q'_r$ ) and closed ( $Q_0'', Q_r''$ ) loop–controllers:

$$Q'_0 = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(\mathcal{R}_0^2 \frac{s_P}{s_N}) d\omega, \quad Q'_r = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(\mathcal{R}_r^2 \frac{s_N}{s_P}) d\omega,$$

$$Q_0'' = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(|W_{ZP}|^2 \frac{s_P}{s_N}) d\omega, \quad Q_r'' = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(|W_{UN}|^2 \frac{s_N}{s_P}) d\omega.$$

Here,  $W_{ZP}$  and  $W_{UN}$  are corresponding complex frequency responses (see details in [1]).

### 1.3. Further Generalizations and the Subject of the Paper

Further studies of this type optimal linear control systems were concerned to a natural generalization for the systems with a non-trivial program ( $Y_0 \neq 0$ ). At the same time, it was an interesting and mathematically natural question to study similar (degenerated) variational problems for various informational functionals similar to (1) with a “power” constraint of type (3).

The first steps in this direction were done by the author (in a collaboration with N. Bukanov and V. Kublanov). We have reported our partial results in the paper [7] which was, in fact, a short conference proceedings announcement. No details of those results were ever published (to the best of my knowledge). It should be mentioned that some results and the intrinsic ideology—to consider an “entropy criterion” of optimal regulation proposed in our paper [7] were rediscovered in 1986 in the announcement [8].

One of the aims of this manuscript is to revise and to generalize the results reported in [7] and to consider them in a general context of variational problems with Legendre–Fenchel dual functionals.

The paper is organized as follows: Section 2 contains a reminder and a generalization of basic results from [1]. We introduce the notations and remind main definitions of all important notions, objects, and formulas used in the manuscript—among them are: the Pinsker functionals of entropy and informational rates of one stationary Gaussian signal with respect to another, signal spectral densities, and transfer functions.

First, we remind readers about (and generalize in the case of a non-trivial Gaussian stationary program signal) the variational problem for the Pinsker functional in the form proposed by Bukanov. We observe and discuss (Proposition 1) an important formula that represents the Pinsker informational rate functional associated with the linear system in Figure 1 in a Shannon-like channel capacity form.

Theorem 2 in Section 3 solves two degenerate variational problems for channel capacities functionals related with the program channel and with the error signal perturbation channel. We compare the expressions for both channel capacities with the Petrov–Uskov Formula (6) from [6]. Theorem 3 (essentially reported in [7]) gives the main balance relations for the system on Figure 1. We want to stress that the observed similarity of these relations with the famous Brillouin and Schrodinger “negentropy” notion appeared due to my numerous discussions with Bukanov, who also indicated a close connection with the information gain, or the Kullback–Leibler divergence (see, for example [9], for a modern geometric approach to “informational thermodynamics”). In Section 4, we consider another informational functionals with or without necessary connections with optimization of linear automatic control systems. Our main observation here is Theorem 4 which remarks that the formula for the second Pinsker main functional—the entropy rate of a random process relatively to another random process—can be interpreted like a Legendre duality relation for the first Pinsker informational rate functional.

## 2. Optimization of the Linear Control System with a Non-Trivial Stationary Gaussian Program

The results of the section (slightly) generalize and extend the material of the paper [1] and give a new interpretation of it. It is worth noting that our notations are different from the original and our motivations and reasonings do not base on possible engineering applications and are purely theoretical.

### 2.1. Useful Formulas and Notations

We collect here (for the convenience of readers) all basic theoretical formulas which we shall use throughout the text. We clarify (once and forever in the paper) that all random variables and random processes (discrete or continuous) we shall understand in the same sense as in the Pinsker book [2], and we refer a “mathematically oriented” reader to this book for precise definitions, description, and properties.

2.1.1. Formulae

In what follows, we shall use the following important formulas. Let  $\zeta, \eta$  be two *random functions* (= stationary (generalized) Gaussian random processes).

- The Gelfand–Pinsker–Yaglom functional (Formula (2.58) from [10]) for the amount of information about a random function contained in another such function:

$$J_{\zeta,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{\zeta}s_{\eta}}{s_{\zeta}s_{\eta} - |s_{\zeta\eta}|^2} d\omega = -\frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(1 - \mathcal{R}_{\zeta\eta}^2) d\omega$$

where  $s_{\zeta}(\omega)$  and  $s_{\eta}(\omega)$  are respectively the spectral densities of  $\zeta$  and  $\eta$  and  $s_{\zeta\eta}(\omega)$ —the mutual spectral density of  $\zeta$  and  $\eta$ . Here,

$$\mathcal{R}_{\zeta\eta}^2 := \frac{|s_{\zeta\eta}|^2}{s_{\zeta}s_{\eta}}$$

is the common correlation function for the processes  $s_{\zeta}$  and  $s_{\eta}$  (Formula (10.21) from [2]).

- Gelfand–Pinsker–Yaglom functional for Gaussian  $\zeta, \eta = \zeta + \nu$  such that  $\zeta$  and  $\nu$  are non-correlated (Formula (2.61) from [10]):

$$J_{\zeta,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log\left(1 + \frac{s_{\nu}}{s_{\zeta}}\right) d\omega$$

- The entropy rate of a random variable  $\zeta$  with respect to  $\eta$  was defined by Pinsker [2]. For two one-dimensional Gaussian stationary (in the wide sense) stochastic processes  $\zeta$  and  $\eta$ , the formula for the entropy rate (Formula (10.5.7) from [2]) reads

$$H_{\zeta,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left(\frac{s_{\zeta}}{s_{\eta}} - 1 - \log \frac{s_{\zeta}}{s_{\eta}}\right) d\omega \tag{9}$$

- For Gaussian  $\zeta, \eta = \zeta + \nu$  such that  $\zeta$  and  $\nu$  are non-correlated:

$$H_{\zeta,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left(\frac{s_{\nu}}{s_{\zeta}} - \log\left(1 + \frac{s_{\nu}}{s_{\zeta}}\right)\right) d\omega$$

2.1.2. Generalized Variational Problem

We suppose now that the stationary Gaussian program  $Y_0 \neq 0$ . Then, the following system of linear operator equations with respect to the variables  $U$  and  $Z$  in  $p$ -coordinates of the Laplace transform for our control system in Figure 1 takes place:

$$\begin{cases} K_r(p)Z(p) - U(p) = K_r(p)N(p), \\ Z(p) + K_0(p)U(p) = K_0(p)P(p) - Y_0(p). \end{cases} \tag{10}$$

The main determinant of the system

$$\Delta = \begin{vmatrix} K_r & -1 \\ 1 & K_0 \end{vmatrix} = 1 + K_0K_r \neq 0.$$

We remind readers about frequency representations  $K_0(j\omega) := \mathcal{R}_0(\omega) \exp(j\phi_0(\omega))$  and  $K_r(j\omega) := \mathcal{R}_r(\omega) \exp(j\phi_r(\omega))$

2.1.3. Notations

For the automatic control system in Figure 1, we introduce the following notations:

1.

$$\mathcal{S}_{0,P,N} := \mathcal{R}_0^2 s_P + s_N$$

2.

$$\mathcal{S}_{r,N,P} := \mathcal{R}_r^2 s_N + s_P$$

3.

$$\mathcal{S}_{0,r,\phi} := (1 + 2\mathcal{R}_0\mathcal{R}_r \cos(\phi_0 + \phi_r) + \mathcal{R}_0^2\mathcal{R}_r^2) s_P s_N$$

In these notations, the functional (2)  $J_{\xi,\eta}$  for the processes  $\xi = U$  and  $\eta = Z$  is written as

$$J_{U,Z} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{\mathcal{S}_{0,P,N}\mathcal{S}_{r,N,P}}{\mathcal{S}_{0,r,\phi}} d\omega \tag{11}$$

Solving the system (10) and calculating the spectral density of signals  $U, Z$  and denoting the program signal  $s_{Y_0}$ , we obtain

$$\begin{cases} s_Z(\omega) = |W_{ZY_0}|^2(j\omega)s_{Y_0}(\omega) + |W_{ZN}|^2(j\omega)s_N(\omega) + |W_{ZP}|^2(j\omega)s_P(\omega), \\ s_U(\omega) = |W_{UY_0}|^2(j\omega)s_{Y_0}(\omega) + |W_{UN}|^2(j\omega)s_N(\omega) + |W_{UP}|^2(j\omega)s_P(\omega). \end{cases} \tag{12}$$

One can find the explicit expressions for the transfer functions  $W(j\omega)$  in Appendix B. We have computed the full functional generalizing functional (2):

$$J_{U,Z}^{gen} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_P[(s_{Y_0} + \mathcal{R}_0^2\mathcal{S}_{r,N,P})(s_{Y_0} + \mathcal{S}_{0,P,N})]}{\mathcal{S}_{0,r,\phi}(s_{Y_0} + \mathcal{R}_0^2s_P)} d\omega \tag{13}$$

We see that  $Y_0 = 0$  (13) and (11) coincide.

**Theorem 1.** *The functional  $J_{U,Z}^{gen}$  has the following extremals:*

$$\begin{cases} \phi_r = \hat{\phi}_r := -\phi_0, \\ \mathcal{R}_r = \hat{\mathcal{R}}_r := \mathcal{R}_0 \frac{s_P}{s_N} [1 + \frac{s_{Y_0}}{\mathcal{R}_0^2 s_P}] \end{cases} \tag{14}$$

**Proof of Theorem 1.** We denote by  $\mathcal{L}$  the ‘‘Lagrangian’’ density for the functional (13):

$$\mathcal{L}(\phi_r(\omega), \mathcal{R}_r(\omega)) = \log \frac{s_P[(s_{Y_0} + \mathcal{R}_0^2\mathcal{S}_{r,N,P})(s_{Y_0} + \mathcal{S}_{0,P,N})]}{\mathcal{S}_{0,r,\phi}(s_{Y_0} + \mathcal{R}_0^2s_P)}$$

The ‘‘degeneracy’’ of (13) means that the Lagrangian density does not depend on derivatives of  $\phi_r$  and  $\mathcal{R}_r$  and the Euler–Lagrange equations are replaced by easy and straightforward computations from the system of variations

$$\begin{cases} \frac{\delta J_{U,Z}^{gen}}{\delta \phi_r} = 0, \\ \frac{\delta J_{U,Z}^{gen}}{\delta \mathcal{R}_r} = 0 \end{cases} \tag{15}$$

(due to the basic lemma of the calculus of variations) by the usual function minimax conditions. The first variation gives

$$\frac{\delta J_{U,Z}^{gen}}{\delta \phi_r} = 0 \Rightarrow -\frac{2\mathcal{R}_0\mathcal{R}_r \sin(\phi_0 + \phi_r)}{\mathcal{S}_{0,r,\phi}(\mathcal{R}_0^2s_P + s_{Y_0})} = 0$$

and we obtain the first condition in Theorem 1. The second condition is verifying with a bit more cumbersome computation and we put it in Appendix A.  $\square$

**Corollary 1.**

1. Taking  $Y_0 = 0$  immediately obtains the extremals (5) of the functional (2);
2. The extremal minimal value of the functional (13) is

$$\widehat{J}_{U,Z}^{gen} = 0,$$

but, exactly like in the case of the zero program (see (15) in [1]), it does not mean that the regulator of the control system does not use the information. It just means that there is a balance of various channel capacities generalizing (8). We shall discuss and interpret these generalizations in the next sections.

3. The same extremals (14) minimize the (generalized) functional of “power” error signal:

$$D_{Z,Z}^{gen} = \int_{-\infty}^{+\infty} \frac{s_{PSN}(s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P})}{\mathcal{S}_{0,r,\phi}} d\omega. \tag{16}$$

4. This extremal value easily computed

$$\widehat{D}_{Z,Z}^{gen} = (\sigma_Z^2)_{min} = \int_{-\infty}^{+\infty} \frac{s_N(s_{Y_0} + \mathcal{R}_0^2 s_P)}{s_{Y_0} + \mathcal{S}_{0,P,N}} d\omega.$$

and coincides with the minimal value of  $D_{Z,Z}$  (14.1) in [1] for zero-program case.

5. The same is true for the similar generalization of the functional (7) for the power of the regulator. We omit the corresponding evident formulas generalizing (14.2) in [1].

We shall obtain another remarkable representation for the functional (13) using the spectral densities of signals  $Z$  and  $\widehat{Z}$ :

**Proposition 1.**

$$J_{U,Z}^{gen} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_Z}{s_{\widehat{Z}}} d\omega \tag{17}$$

**Proof.** We shall compute the spectral density  $s_Z$  from the first equation of (12):

$$s_Z(\omega) = \frac{(s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P}) s_{PSN}}{\mathcal{S}_{0,r,\phi}},$$

the spectral density  $s_{\widehat{Z}}$  putting the extremals (14) in the above formula and substituting both spectral densities in (17). □

This formula will be very useful in further interpretations in Section 4.

**3. Capacity of Channels and Balance Theorems for the Linear Stationary Control System**

Now, using the generalization for  $Y_0 \neq 0$ , we can also generalize and extend the relations which we have reported in [7]. In the same time, we shall propose some new interpretations of these relations. We shall assume throughout this section that all signals are stationary and Gaussian.

*3.1. Capacity of Channels via Differential Entropy*

We recall that, for the capacity of the channel for the linear control system in Figure 1, the given program  $Y_0$  can be computed via differential entropies (“reduced to a degree of freedom”) of this program and the error (defect)  $Z$  (see, for example, (ch. IV (12.5) in [11]):

$$\mathcal{C}_{Y_0} = \mathcal{H}(Y_0) - \mathcal{H}(Z) \tag{18}$$

where

$$\begin{cases} \mathcal{H}(Y_0) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(2\pi e) s_{Y_0}(\omega) d\omega, \\ \mathcal{H}(Z) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(2\pi e) s_Z(\omega) d\omega \end{cases} \quad (19)$$

and the expression (18) is written as

$$C_{Y_0} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0}(\omega)}{s_Z(\omega)} d\omega. \quad (20)$$

Analogously, one can obtain the capacity of the channel for the linear control system relatively the perturbations  $P$  and  $N$ :

$$C_{P,N} = \mathcal{H}(Z) - \mathcal{H}(P, N) \quad (21)$$

and

$$\mathcal{H}(P, N) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(2\pi e) \mathcal{S}_{0,P,N}(\omega) d\omega$$

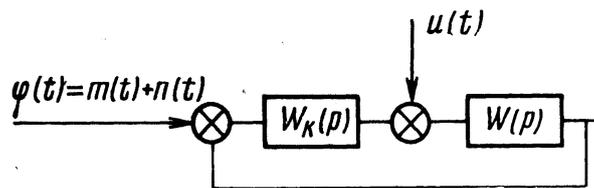
such that

$$C_{P,N} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_Z(\omega)}{\mathcal{S}_{0,P,N}(\omega)} d\omega. \quad (22)$$

**Remark 2.** It would be interesting to compare Formulas (20) and (22) with Formula (16) from [6]:

$$C = \int_W \log \frac{s_m(\omega)}{s_\epsilon(\omega)} d\omega. \quad (23)$$

for a channel capacity of the linear system in Figure 2.



**Figure 2.** This is a figure from [5]. Here,  $W(p)$  denotes the transfer function of the operator object. The signal  $m(t)$  with the perturbation signal  $n(t)$  form a function  $\phi(t)$  on the input of a regulator given by an operator  $W_k(p)$ . The system should reproduce the “useful” signal  $m(t)$  as an output.

**Remark 3.** Here the “cross” notation of the diagram nodes means the same as the “addition” nodes notation in Figure 1.

Here,  $s_m(\omega)$  and  $s_\epsilon$  are spectral densities of the “useful” signal  $m(t)$  and the “error” signal  $\epsilon(t)$  correspondingly.

The integral should be calculated over the total frequency band  $W$ . In particular, if  $u(t) = 0$  and  $m$  and  $n$  are non-correlated, then

$$C = \int_W \log \frac{s_m(\omega) + s_n(\omega)}{s_n(\omega)} d\omega. \quad (24)$$

We assume that the perturbation  $n(t)$  has the spectral density with an additional “white noise” spectral density:  $s_n(\omega) = \tilde{s}_n(\omega) + c^2$ . We have from (24)

$$C = \int_W \log \frac{s_m(\omega) + \tilde{s}_n(\omega) + c^2}{\tilde{s}_n(\omega) + c^2} d\omega. \quad (25)$$

The analog of the the balance relations (18) and (21) reads

$$\mathcal{H}(n, m) - \mathcal{H}(n, \epsilon) = \mathcal{C} = \frac{1}{W} \int_W \log \frac{s_m(\omega) + \tilde{s}_n(\omega) + c^2}{\tilde{s}_n(\omega) + c^2} d\omega.$$

The authors of [5] argue that the entropy quantity  $\mathcal{H}(n, \epsilon)$  is a characteristic of information lost while the signal goes through the system of Figure 2, and the minimum of this loss is achieved when  $c \equiv 0$ . When  $c \rightarrow \infty$ , the entropy  $\mathcal{H}(n, \epsilon) \rightarrow \mathcal{H}(n, m)$  i.e., the “error” entropy tends to the entropy of the “useful” signal and the channel capacity  $\mathcal{C}$  goes to 0.

**Remark 4.** We should stress here that:

1. all signals are supposed to be stationary random Gaussian with zero expectation value;
2. the “optimality” solution in [5] was founded with the type “energy” constraint (3), which means in the context the minimality condition of the second momentum (“dispersion”) for the error signal  $s_\epsilon$  :

$$\int_W s_\epsilon d\omega = \int_W \frac{s_m(\tilde{s}_n + c^2)}{s_m + \tilde{s}_n + c^2} d\omega = \sigma_\epsilon^2$$

3. The minimal value  $\min \sigma_\epsilon^2$  is achieved when  $c = 0$  and this condition shows that the “energy” and informational criteria give the same result (compare with the similar conclusion for the case of the system in Figure 1)
4. The independency condition (the signals  $m(t)$  and  $n(t)$  are non-correlated) can be released, and this leads to the following generalization of capacity channel formula (24):

$$\mathcal{C} = \int_W \log \frac{s_m(s_m + s_n)}{s_m s_n + |s_{mn}|^2} d\omega. \tag{26}$$

### 3.2. Variational Problems for the Channel Capacity Functionals

Now, we need to substitute the spectral densities of all signals in Formulas (20) and (22) and, as a result, obtain the following two channel capacity functionals:

$$\mathcal{C}_{Y_0} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0} \mathcal{S}_{0,r,\phi}}{s_P s_N (s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P})} d\omega \tag{27}$$

and

$$\mathcal{C}_{P,N} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_P s_N (s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P})}{\mathcal{S}_{0,P,N} \mathcal{S}_{0,r,\phi}} d\omega \tag{28}$$

**Remark 5.** It is evident that the functional (28) generalizes the similar thing for the  $Y_0 = 0$  case

$$\mathcal{C}_{P,N}|_{Y_0=0} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_P s_N \mathcal{R}_0^2 \mathcal{S}_{r,N,P}}{\mathcal{S}_{0,P,N} \mathcal{S}_{0,r,\phi}} d\omega \tag{29}$$

whose extremum under conditions (5) are

$$\mathcal{C}_{P,N}^{min}|_{Y_0=0} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_P s_N \mathcal{R}_0^2}{\mathcal{S}_{0,P,N}^2} d\omega \tag{30}$$

**Remark 6.** (One can straightforwardly check that this extremal value is a sum of the informational performances ( $Q''_0, Q''_r$ ) of the object and the regulator channels with closed loop–controllers outputs under the conditions (5):

$$\mathcal{C}_{P,N}^{min,0}|_{Y_0=0} = (Q''_0 + Q''_r)|_{\phi=\phi_0, \mathcal{R}_r=\mathcal{R}_0 \frac{s_P}{s_N}}$$

These remarks definitely indicate that there are more general information–performance balance relations and we start to discuss them.

**Theorem 2.** *The functional (27) has the maximal value*

$$C_{Y_0}^{max} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0}}{s_N} \left[ 1 + \frac{s_N}{s_{Y_0} + \mathcal{R}_0^2 s_P} \right] d\omega \tag{31}$$

and the functional (28) has the minimal value

$$C_{P,N}^{min} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{(s_{Y_0} + \mathcal{R}_0^2 s_P) s_N}{(s_{Y_0} + \mathcal{S}_{0,P,N}) \mathcal{S}_{0,P,N}} d\omega \tag{32}$$

for the same set of extremals (14):

$$\begin{cases} \phi_r = \hat{\phi}_r := -\phi_0, \\ \mathcal{R}_r = \hat{\mathcal{R}}_r := \mathcal{R}_0 \frac{s_P}{s_N} \left[ 1 + \frac{s_{Y_0}}{\mathcal{R}_0^2 s_P} \right] \end{cases}$$

**Proof of Theorem 2.** Both functionals are degenerated in the same sense as above. Therefore, their extremals are straightforwardly reduced to their Lagrangian densities’ usual minimax computations. Similarly, to determine a character of obtained extremums, it is enough to verify the Hessian matrices definiteness for these densities instead of computations of the second variation  $\delta^2(C_{Y_0})$  and  $\delta^2(C_{N,P})$ . We omit these routine but tedious computations which are absolutely similar to those in Theorem 1.  $\square$

### 3.3. Informational Balance in the Linear Control System

The following theorem was obtained in a collaboration with Bukanov and Kublanov and was announced in [7].

**Theorem 3.** *There are two balance relations between the channel capacities:*

1. *For the channel capacity of the program  $Y_0$  and the informational rate in the control signal  $U$  in the “defect” signal  $Z$ :*

$$C_{Y_0}^{max} - C_{Y_0} = J_{U,Z}^{gen}. \tag{33}$$

2. *For the channel capacity of perturbations  $P$  and  $N$  and the above informational rate:*

$$C_{P,N} - C_{P,N}^{min} = J_{U,Z}^{gen}. \tag{34}$$

**Proof.** Both identities are easily checked by manipulations with definitions of the corresponding channel capacities and computed extrema. Let us verify the first one:

$$\begin{aligned} C_{Y_0}^{max} - C_{Y_0} &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0}}{s_N} \left[ 1 + \frac{s_N}{s_{Y_0} + \mathcal{R}_0^2 s_P(\cdot)} \right] d\omega - \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0} \mathcal{S}_{0,r,\phi}}{s_P s_N (s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P})} d\omega \\ &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_P (s_{Y_0} + \mathcal{R}_0^2 s_P + s_N) (s_{Y_0} + \mathcal{R}_0^2 \mathcal{S}_{r,N,P})}{\mathcal{S}_{0,r,\phi} (s_{Y_0} + \mathcal{R}_0^2 s_P)} d\omega = J_{U,Z}^{gen}. \end{aligned}$$

$\square$

**Remark 7.** *Using (20), one can recast Formula (17) re-writing the first balance relation (33) as*

$$J_{U,Z}^{gen} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0}(\omega)}{s_{\hat{Z}}(\omega)} d\omega - \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{Y_0}(\omega)}{s_Z(\omega)} d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_Z(\omega)}{s_{\hat{Z}}(\omega)} d\omega.$$

On other hand, it means that

$$J_{U,Z}^{gen} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(2\pi e) s_Z(\omega) d\omega - \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log(2\pi e) s_{\hat{Z}}(\omega) d\omega = \mathcal{H}(Z) - \mathcal{H}(\hat{Z})$$

which shows a similarity with the Brillouin “negentropy” principle which says that the system loses information during a transition from a state with low entropy to a state with higher entropy. (We include a brief description of this notion in the Discussion section below).

Naively speaking, the entropy  $\mathcal{H}(Z)$  of the error signal of the linear system with any regulator is bigger (“more chaos”) than the error signal entropy  $\mathcal{H}(\hat{Z})$  for the case of the system with optimal regulator:

$$\mathcal{H}(\hat{Z}) = \mathcal{H}(Z) - J_{U,Z}^{gen} \geq 0.$$

Similarly, the same result can be obtained from the second balance relation

$$C_{P,N} - C_{P,N}^{min} = \mathcal{H}(Z) - \mathcal{H}(P,N) - \mathcal{H}(\hat{Z}) + \mathcal{H}(P,N) = J_{U,Z}^{gen}.$$

One can imagine here the informational rate  $J_{U,Z}^{gen}$  as an “informational difference” between  $Z$  and  $\hat{Z}$  or as a “Brillouin–Schrödinger negentropy” [12].

**Remark 8.** We should admit that the first balance relation (33) does not exist in the absence of the program signal ( $Y_0 = 0$ ), while the second is a generalization of the previously described informational balance relation (8): The LHS of the identity (34) under condition  $Y_0 = 0$  is:

$$\begin{aligned} (C_{P,N} - C_{P,N}^{min})|_{Y_0=0} &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{P^S N} \mathcal{R}_0^2 s_{r,N,P}}{S_{0,P,N} S_{0,r,\phi}} d\omega - \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{s_{P^S N} \mathcal{R}_0^2}{S_{0,P,N}^2} d\omega \\ &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{S_{0,P,N} S_{r,N,P}}{S_{0,r,\phi}} d\omega = J_{U,Z} = (J_{U,Z}^{gen})|_{Y_0=0}. \end{aligned} \tag{35}$$

which is exactly the RHS of (34) in the same condition.

#### 4. Entropy Rate Functional for the Linear Stationary Control System

Now, we shall discuss the entropy rate functional (9) for the pair of Gaussian stationary error signal ( $Z$ ) and the error signal with optimal regulator ( $\hat{Z}$ ) in the linear control system above:

$$H_{Z,\hat{Z}} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left( \frac{s_Z}{s_{\hat{Z}}} - 1 - \log \frac{s_Z}{s_{\hat{Z}}} \right) d\omega. \tag{36}$$

Introduce the function  $F(X) := \log(X) - 1 + \frac{1}{X}$  of a positive argument  $X > 0$  and study its behavior.

##### 4.1. Properties of the Function $F(X)$ and Legendre–Fenchel Transformation

**Lemma 1.** The function  $F(X)$  has the following (almost) evident properties:

1.  $F(X) \geq 0$  and  $F(1) = 0$ .
2. Let  $X = \frac{u}{v}$  such that  $u > 0; v > 0$ , then

$$\log\left(\frac{u}{v}\right) \geq \left(1 - \frac{v}{u}\right);$$

3. If  $u = u(\omega)$  and  $v = v(\omega)$  are integrable on  $(\mathbb{R}, d\mu(\omega))$  with some measure  $\mu(\omega)$  and

$$\int_{\mathbb{R}, \mu} u d\mu(\omega) = \int_{\mathbb{R}, \mu} v d\mu(\omega)$$

then

$$\int_{\mathbb{R}, \mu} u \log\left(\frac{u}{v}\right) d\mu(\omega) \geq 0;$$

4. Let  $Y = \frac{1}{X}$  and  $\mathcal{F}(Y) := F\left(\frac{1}{X}\right)$ ; then, the function  $\mathcal{F}(Y) = Y - 1 - \log(Y)$  admits the Legendre–Fenchel transformation  $\mathcal{F}^*(p) = -\log(1 - p)$ .

**Proof.**

1. To prove the first positivity property of  $F(X)$ , we split the half line  $\{X \geq 0\}$  in two subsets:  $\{0 \leq X < 1\}$  and  $\{X \geq 1\}$ . Start with the second subset: for  $1 \leq t \leq X$ , we have  $\frac{1}{t} \geq \frac{1}{X}$  and

$$\log X = \int_1^X \frac{dt}{t} \geq \int_1^X \frac{dt}{X} = \frac{1}{X} \int_1^X dt = \frac{X-1}{X} = 1 - \frac{1}{X}.$$

In the first subset, one has similarly  $-\frac{1}{t} \geq -\frac{1}{X}$  if  $X \leq t < 1$  and

$$\log X = -\int_X^1 \frac{dt}{t} \geq -\int_X^1 \frac{dt}{X} = 1 - \frac{1}{X}.$$

2. Tautological corollary of the computations in (1).  
 3. Using the inequality (2), we obtain:

$$\int_{\mathbb{R}, \mu} u \log\left(\frac{u}{v}\right) d\mu(\omega) \geq \int_{\mathbb{R}, \mu} u\left(1 - \frac{v}{u}\right) d\mu(\omega) = 0;$$

4. The Legendre–Fenchel transformation  $\mathcal{F}^*(p)$  of the function  $\mathcal{F}(Y)$  exists because the function is smooth and convex for  $Y > 0$ . The only critical point  $Y_{crit} = 1$  with the critical value  $\mathcal{F}(Y_{crit}) = 0$  is obtained from  $1 - \frac{1}{Y} = 0$ . By the definition (see, f.e. [13], we put  $p := 1 - \frac{1}{Y}$  and find  $Y(p) = \frac{1}{1-p}$ . Then, we find the Legendre-dual function from  $\frac{d}{dp}(\mathcal{F}^*(p)) = \frac{1}{1-p}$ ,  $\mathcal{F}^*(p) = -\log(1 - p)$ . We shall check that it satisfies the duality relation

$$\mathcal{F}^*(p) = pY(p) - \mathcal{F}(Y(p)) :$$

$$pY(p) - \mathcal{F}(Y(p)) = p\frac{1}{1-p} - \left(\frac{1}{1-p} - 1 - \log\left(\frac{1}{1-p}\right)\right) = -\log(1 - p) = \mathcal{F}^*(p).$$

□

We note the famous Fenchel–Young inequality which reads in this case as

$$\mathcal{F}(Y) + \mathcal{F}^*(p) \geq Yp$$

which gives the evident inequality

$$1 + \log[Y(1 - p)] \leq Y(1 - p).$$

#### 4.2. Legendre Transformation Analogy for the Control System Functionals

We can observe that the Pinsker entropy rate functional (9) can be written in the form

$$H_{\xi,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left( \frac{s_{\xi}}{s_{\eta}} - 1 - \log \frac{s_{\xi}}{s_{\eta}} \right) d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \mathcal{F}(Y(\omega)) d\omega,$$

where  $Y(\omega) := \frac{s_{\xi}(\omega)}{s_{\eta}(\omega)}$ . Then, the Legendre–Fenchel dual functional should be

$$\frac{1}{4\pi} \int_{-\infty}^{+\infty} \mathcal{F}^*(p(\omega)) d\omega,$$

with  $p(\omega) = 1 - \frac{1}{Y(\omega)} = 1 - \frac{s_{\eta}(\omega)}{s_{\xi}(\omega)}$ . Using the part 4 of the Lemma 1, one can compute

$$\mathcal{F}^*(p(\omega)) = -\log(1 - p(\omega)) = -\log\left(\frac{s_{\eta}(\omega)}{s_{\xi}(\omega)}\right) = \log\left(\frac{s_{\xi}(\omega)}{s_{\eta}(\omega)}\right) = \log Y(\omega).$$

Thus, we have obtained the following Legendre duality relation:

$$\frac{1}{4\pi} \int_{-\infty}^{+\infty} \mathcal{F}(Y(\omega)) d\omega + \frac{1}{4\pi} \int_{-\infty}^{+\infty} \mathcal{F}^*(p(\omega)) d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} p(\omega) Y(\omega) d\omega.$$

Coming back to the Pinsker entropy rate functional Formula (9), one observes that the duality implies

$$H_{\xi,\eta} + \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log\left(\frac{s_{\xi}(\omega)}{s_{\eta}(\omega)}\right) d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left(\frac{s_{\xi}(\omega)}{s_{\eta}(\omega)} - 1\right) d\omega. \tag{37}$$

Formula (37) expressed the duality relation between the entropy rate and the functional  $\frac{1}{4\pi} \int_{-\infty}^{+\infty} \log\left(\frac{s_{\xi}(\omega)}{s_{\eta}(\omega)}\right) d\omega$  for any pair of stationary Gaussian signals. Let us consider the case of the Pinsker entropy rate for our control system (36), i.e.,  $s_{\xi}(\omega) = s_Z(\omega)$  and  $s_{\eta}(\omega) = s_{\hat{Z}}(\omega)$ . In this case, (37) and (36) give us the Legendre duality condition between the functionals:

$$H_{Z,\hat{Z}} + \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log\left(\frac{s_Z(\omega)}{s_{\hat{Z}}(\omega)}\right) d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left(\frac{s_Z(\omega)}{s_{\hat{Z}}(\omega)} - 1\right) d\omega. \tag{38}$$

Summing up the discussion in Section 4.2, one can formulate the following theorem:

**Theorem 4.** *The Pinsker functional of the entropy rate of the error signal in the stationary linear control system with respect to the error signal of this control system with the optimal by the informational criterion regulator and the informational rate functional for these signals are in Legendre duality:*

$$H_{Z,\hat{Z}} + J_{U,Z}^{gen} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left(\frac{s_Z(\omega)}{s_{\hat{Z}}(\omega)} - 1\right) d\omega. \tag{39}$$

**Remark 9.** *We can also check that the the duality relation (37) can be easily generalized to the multidimensional case for two  $n$ – dimensional stationary Gaussian processes  $\xi$  and  $\eta$  such that  $\xi = (\xi_1, \dots, \xi_n), \eta = (\eta_1, \dots, \eta_n)$  and such that the pairs  $(\xi_j, \eta_j)$  are pair-wise uncorrelated. Then, one can define  $n$ – dimensional analogues of entropy rate and informational rate functionals:*

$$H_{\xi,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left[ \left( \sum_{j,k=1}^n s_{\xi_j \xi_k}(\omega) s_{\eta_j \eta_k}^{-1}(\omega) \right) - n - \log \frac{\det \|s_{\xi_j \xi_k}(\omega)\|}{\det \|s_{\eta_j \eta_k}(\omega)\|} \right] d\omega$$

which can be simplified (because of the mutual independence of the signals  $\xi$  and  $\eta$ ) to

$$H_{\xi,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \sum_{j=1}^n \left( \frac{s_{\xi_j \xi_j}(\omega)}{s_{\eta_j \eta_j}(\omega)} - 1 - \log \frac{s_{\xi_j \xi_j}(\omega)}{s_{\eta_j \eta_j}(\omega)} \right) d\omega, \tag{40}$$

where  $\|s_{\xi_j \xi_k}(\omega)\|$  (resp.  $\|s_{\eta_j \eta_k}(\omega)\|$ ) is the mutual spectral densities matrix of signal-components of  $\xi$  (resp.  $\eta$ ) and  $s_{\eta_j \eta_k}^{-1}(\omega)$  denotes the  $(jk)$  entry of the matrix  $\|s_{\eta_j \eta_k}(\omega)\|$ .

We see that one can apply the Legendre duality Formula (37) to each summand of (40) and re-write it as

$$H_{\xi,\eta} + L_{\xi,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \sum_{j=1}^n \left( \frac{s_{\xi_j \xi_j}(\omega)}{s_{\eta_j \eta_j}(\omega)} - 1 \right) d\omega, \tag{41}$$

where the functional

$$L_{\xi,\eta} = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \frac{\det \|s_{\xi_j \xi_k}(\omega)\|}{\det \|s_{\eta_j \eta_k}(\omega)\|} d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \sum_{j=1}^n \log \frac{s_{\xi_j \xi_j}(\omega)}{s_{\eta_j \eta_j}(\omega)} d\omega$$

plays the role of the Pinsker informational rate functional (see chapters 10 and 11 in [2]) and for  $n = 1$  coincides with  $\frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \left( \frac{s_{\xi}(\omega)}{s_{\eta}(\omega)} \right) d\omega$ .

We want to stress that our multidimensional considerations are purely formal. We do not discuss the conditions of existence and finiteness for all formulas above and refer to [2,14] for all necessary mathematical details. We also do not know if these multidimensional generalizations can immediately apply to an optimal automatic control system problem, but we hope that such functional might be useful in comparison with the method of multidimensional optimization where the minimum of the error expectation value square and maximum of the information mean value criteria were used [15].

### 5. Discussion

The relationship between entropy and information was first discovered in the seminal work of Szilard [16]. Later, in the works of Brillouin [17], the negentropy principle of information was formulated, generalizing the second law of thermodynamics. According to this principle, both entropy and information should be considered jointly and cannot be interpreted separately. Brillouin had shown [18] that the definition of “information” leads to a direct connection between information and the negative of entropy (here is the abbreviation “negentropy”). Every experiment consumes negentropy (increases entropy) and yields information. The negentropy principle of information is a generalization of Carnot’s principle in thermodynamics that Brillouin had formulated in the following form: the amount of information contained in a physical system must be considered as a negative term in the total entropy of this system. If a system is isolated, it fulfills Carnot’s principle in its generalized form. According to this principle, the total entropy of the system does not decrease. Our Theorem 4 is a good illustration of these energy-information principle.

The functional  $L_{\xi,\eta} := \frac{1}{4\pi} \int_{-\infty}^{+\infty} \log \left( \frac{s_{\xi}(\omega)}{s_{\eta}(\omega)} \right) d\omega$  enters in the Legendre duality relation (37) with the entropy rate functional  $H_{\xi,\eta}$  and has explicit interpretation in terms of the informational rate  $J_{Z,\hat{Z}}$  in the case of the control system. The Legendre balance relation predicts some thermodynamical allusions going back to Gibbs thermodynamical potential relation:

$$S(E) + F(\beta) = \beta E,$$

which relates the Helmholtz free energy  $F$  to the entropy  $S$ , total energy  $E$ , and  $\beta = \frac{1}{T}$  is inverse to the temperature (the Boltzman constant is supposed to be 1 here). If we shall follow the initial idea of Bukanov about a “similarity” of the informational rate functional  $L_{\xi,\eta}$  with the Brillouin negentropy

$\mathcal{N}(\rho)$  and the information gain or Kullback–Leibler divergence  $I(p, q) = \int \rho \log \rho dq$  which is “similar (up to sign) to the differential entropy expression”:

$$\mathcal{N}(\rho) = I(\xi) + I(p, q) = - \int \xi \log \xi dq + I(p, q) = \int \log \frac{\rho}{\xi} dq$$

for Gaussian  $\xi$ , which has the same mean value as  $\rho$ . (notations and the statement from p. 15 of [9]), then we should conclude that, in our case, the informational rate functional  $J_{\xi, \eta}$  is rather a Legendre dual to the entropy rate  $H_{\xi, \eta}$ , though it plays a “role of Hamiltonian” while  $H_{\xi, \eta}$  figured here as a “Lagrangian”. Following the “mechanical and thermodynamical” allusions of [9], we could interpret the results of our Theorems 1 and 2 as a (specified) example of the “principle of minimal gain information” and the balance relations in Theorem 3 show the connection with Jaynes principle of maximum entropy. It would interesting to give further interpretations in this line and to understand what the role of the “partition function”  $\mathcal{Z}$  is, which, in our context, should naively read

$$L_{\xi, \eta} = \int \log \frac{s_{\xi}}{s_{\eta}} d\omega = - \int \log \mathcal{Z} d\omega.$$

This question, together with more serious comprehension of appropriate Lagrangian or canonical coordinates in these application approaches, would open a door to powerful invariant methods in the theory of information control systems.

There are many other possible developments to study, but all of them look more traditional. We could almost straightforwardly generalize our results using slightly more complicated cases of the capacity channel formulas in comparison with the results of Petrov et al. for signals with correlations, multidimensional systems generalizing the system in Figure 2, etc. The “physical realization” conditions are also beyond the scope of this paper.

**Funding:** I was partially supported during this work by the Interdisciplinary CNRS Project “Granum” (“Graded Geometry, Numerical methods, Mechanics, and Control Theory”) and Russian Foundation for Basic Research for a partial support under the Grant RFBR 18-01-00461.

**Acknowledgments:** It is my pleasure to dedicate this paper to Nikolay Petrovich Bukanov, who had the courage many years ago to hire me as a young researcher and the patience to introduce me in the subject of applied Information Theory. His ideas and vision were very influential and his support was crucial at some difficult moments at the start of my scientific career. I am also very grateful to V. S. Kublanov for his collaboration on [7] and to V.V. Lychagin for numerous discussions of various aspects of geometric approach to thermodynamics and the allusions from Information theory. He encouraged me to revise my old computations and to transform them in these notes. It is also my pleasure to thank O.N. Poberejnaia for numerous useful discussions in the early stages of this work. Finally, I am very grateful to the anonymous referee No. 2 for his useful remarks.

**Conflicts of Interest:** The author declares no conflict of interest.

### Appendix A. Proof of the Theorem 1, Second Condition

$$\begin{aligned} \frac{\delta J_{U,Z}^{gen}}{\delta \mathcal{R}_r} &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \frac{\delta}{\delta \mathcal{R}_r} \left[ \log(s_{Y_0} + \mathcal{R}_0^2 s_P + \mathcal{R}_0^2 \mathcal{R}_r^2 s_N) - \log \mathcal{S}_{0,r,\phi} \right] d\omega \\ &= \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left[ \frac{s_N \mathcal{R}_r \mathcal{R}_0}{(s_{Y_0} + \mathcal{R}_0^2 s_P + \mathcal{R}_0^2 \mathcal{R}_r^2 s_N)} - \frac{\cos(\phi_0 + \phi_r) + \mathcal{R}_0 \mathcal{R}_r}{1 + 2\mathcal{R}_0 \mathcal{R}_r \cos(\phi_0 + \phi_r) + \mathcal{R}_0^2 \mathcal{R}_r^2} \right] d\omega. \end{aligned}$$

Using the first condition of (14), one can re-write it as

$$\begin{aligned} &\frac{1}{4\pi} \int_{-\infty}^{+\infty} \left[ \frac{s_N \mathcal{R}_r \mathcal{R}_0}{(s_{Y_0} + \mathcal{R}_0^2 s_P + \mathcal{R}_0^2 \mathcal{R}_r^2 s_N)} - \frac{1 + \mathcal{R}_0 \mathcal{R}_r}{1 + 2\mathcal{R}_0 \mathcal{R}_r + \mathcal{R}_0^2 \mathcal{R}_r^2} \right] d\omega = \\ &\frac{1}{4\pi} \int_{-\infty}^{+\infty} \left[ \frac{s_N \mathcal{R}_r \mathcal{R}_0}{(s_{Y_0} + \mathcal{R}_0^2 s_P + \mathcal{R}_0^2 \mathcal{R}_r^2 s_N)} - \frac{1}{1 + \mathcal{R}_0 \mathcal{R}_r} \right] d\omega = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \left[ \frac{s_N \mathcal{R}_r \mathcal{R}_0 - (s_{Y_0} + \mathcal{R}_0^2 s_P)}{(s_{Y_0} + \mathcal{R}_0^2 s_P + \mathcal{R}_0^2 \mathcal{R}_r^2 s_N)(1 + \mathcal{R}_0 \mathcal{R}_r)} \right] d\omega. \end{aligned}$$

Now, using the basic lemma of variational calculus, we obtain the second condition in (14):

$$s_N \mathcal{R}_r \mathcal{R}_0 = s_{Y_0} + \mathcal{R}_0^2 s_P.$$

## Appendix B. Formulas for Transition Functions in Frequency Variables

We shall display for completeness reasons the formulas (from [1,7]) for transition functions of the linear system in Figure 1 which we have used in computations of the spectral densities  $s_U$  and  $s_Z$  (12):

$$\begin{cases} W_{ZY_0}(j\omega) = \frac{-1}{1+\mathcal{R}_0\mathcal{R}_r}, & W_{ZN}(j\omega) = \frac{-\mathcal{R}_0\mathcal{R}_r}{1+\mathcal{R}_0\mathcal{R}_r} \\ W_{ZP}(j\omega) = \frac{\mathcal{R}_0}{1+\mathcal{R}_0\mathcal{R}_r}, & W_{UY_0}(j\omega) = \frac{-\mathcal{R}_r}{1+\mathcal{R}_0\mathcal{R}_r} \\ W_{UN}(j\omega) = \frac{\mathcal{R}_r}{1+\mathcal{R}_0\mathcal{R}_r}, & W_{UP}(j\omega) = \frac{\mathcal{R}_0\mathcal{R}_r}{1+\mathcal{R}_0\mathcal{R}_r}. \end{cases}$$

## References

1. Bukanov, N.P. Information criterion of optimality of automatic control systems. *Avtomat. Telemekh.* **1972**, *1*, 57–62. (In Russian)
2. Pinsker, M.S. *Information and Informational Stability of Random Variables and Processes; Problemy Peredači Informacii*; Izdat. Akad. Nauk SSSR: Moscow, Russia, 1960; Volume 7, 203p. (In Russian); English Translation: Pinsker, M.S., Translated and edited by Amiel Feinstein; Holden-Day, Inc.: San Francisco, CA, USA; London, UK; Amsterdam, The Netherlands, 1964; 243p.
3. Shannon, C. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656. [[CrossRef](#)]
4. Petrov, B.N.; Petrov, V.V.; Ulanov, G.M.; Zaporczets, A.W.; Uskov, A.S.; Kotchubievsky, I.L. *Stochastic Process in Control and Information Systems*, Survey Paper 63; IFAC: Warszawa, Poland, 1969.
5. Petrov, V.V.; Uskov, A.S. Informational aspects of the regularization problem. *Doklady AN SSSR* **1970**, *195*, 780–782. (In Russian)
6. Petrov, V.V.; Uskov, A.S. Channel capacities of feed-back informational systems with internal perturbations. *Doklady AN SSSR* **1970**, *194*, 1029–1031. (In Russian)
7. Bukanov, N.P.; Kublanov, V.S.; Rubtsov, V.N. About a question on a synthesis of an automatic control system optimal with respect to an informational criterion. In Proceedings of the All-Union Conference Optimization and Efficiency of Systems and Processes in Civil Aviation, MIIGA, Moscow, Russia, March 1979; pp. 83–88. (In Russian)
8. Petrov, V.V.; Sobolev, V.I. Entropy approach to a quality criterial of automatic control systems. *Doklady AN SSSR* **1986**, *287*, 799–801. (In Russian)
9. Lychagin, V.V. Contact Geometry, Measurement, and Thermodynamics. In *Nonlinear PDEs, Their Geometry, and Applications*; Kycia, R.A., Ulan, M., Schneider, E., Eds.; Birkhäuser: Basel, Switzerland, 2019; pp. 3–52.
10. Gelfand, I.M.; Yaglom, A.M. Computation of the amount of information about a stochastic function contained in another such function. *Uspehi Mat. Nauk* **1957**, *12*, 3–52. (In Russian)
11. Tarasenko, F.P. *Introduction in the course of Information Theory*; Tomsk State University: Tomsk, Russia, 1963; p. 239. (In Russian)
12. Schrödinger, E. *What Does Life Is from the Physics Point of View?* State Edition of Foreign Literature: Moscow, Russia, 1947; p. 147. (In Russian)
13. Arnold, V.I. *Mathematical Methods of Classical Mechanics*; Translated by K. Vogtmann and A. Weinstein; Graduate Texts in Mathematics, 60; Springer: New York, NY, USA, 1989; 516p.
14. Pinsker, M.S. Entropy, entropy rate and entropy stability of Gaussian random variables and processes. *Doklady AN SSSR* **1960**, *133*, 531–534. (In Russian)
15. Petrov, V.V.; Zaporozhets, A.V. Principle of developable systems. *Doklady AN SSSR* **1975**, *224*, 779–782. (In Russian)
16. Szilard, L. On the reduction of entropy in a thermodynamic system by the intervention of intelligent beings. *Z. Phys.* **1929**, *53*, 840–856. (In German) [[CrossRef](#)]

17. Brillouin, L. *Science and Information Theory*; State Edition of Foreign Literature: Moscow, Russia, 1960; p. 392. (In Russian)
18. Brillouin, L. Thermodynamics, statistics, and information. *Am. J. Phys.* **1961**, *29*, 318–328. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).