V.A. ZHMUD, L.V. DIMITROV, J. NOSEK

NUMERICAL OPTIMIZATION OF REGULATORS FOR AUTOMATIC CONTROL SYSTEM

NOVOSIBIRSK
2019

UDC 621.375.087.9:681.515/516 (075.8)
Z 62

Reviewers:
Dr. of Technical Sciences, Prof. *G.A. Frantsuzova*,
Novosibirsk State Technical University, Novosibirsk, Russia

Dr. of Physics and Mathematical Sciences, Corresponding member of RAN,
Prof. *A.V. Taichenachev*, Institute of Laser Physics, Novosibirsk, Russia

Zhmud Vadim A.

Z 62 Numerical Optimization of Regulators for Automatic Control System: Textbook for higher education / Vadim A. Zhmud, Lubomir V. Dimitrov, Jaroslav Nosek. – Novosibirsk: NSTU Publisher, 2019. – 296 p.

ISBN 978-5-7782-3802-2

The book is intended for students and PhD students in the field of automatics, mechatronics, robotics, digital or analog feedback control. Particullary, in NSTU (Novosibirsk, Russia) it is intended for the students enrolled in the educational field 27.03.04 “Control in Technical Systems” in the course “Computer-aided design tools and control systems” (for Bachelors in their 4th year). It includes training materials and guidelines for self-assessment (questions for self-testing).

The basic knowledge of previously studied mathematical disciplines is required for successful study of this course.

The book is prepared due to the financial support from the European Foundation Erasmus+, projects KA-107A and SmartCity.

UDC 621.375.087.9:681.515/516 (075.8)

ISBN 978-5-7782-3802-2

Content

INTRODUCTORY WORDS

This fascinating manual introduces the reader to the complicated matter of automatic control systems in a pleasant and intriguing manner. I as a control theory specialist read in one breath. It is written in an accessible language and is illustrated with many examples. Moreover, it contains tasks taken from the practice. The textbook is written by authors from Russia, Bulgaria, and Czech Republic and is a result of a joint Tempus Project No. 517138 – JPCR that aims at double degree diplomas on Mechatronics issued by Novosibirsk Technical University and Sofia Technical University or Liberec University. My expertise in theory and practice of Automatic Control Systems makes me to believe that this textbook is useful both for Bachelors and Masters of the Technical University of Sofia who are trained in the fields "Mechatronics," "Mechatronic Systems", and "Industrial Engineering", and in particular in the following disciplines: Electronic control and control devices and systems, Mechatronics systems and management systems, Mechatronic systems and management systems, Control systems and control systems. In addition, this textbook could be of benefit to PhD students and researchers working in areas close to the automatic control system.

Assoc. prof. *Pancho Tomov*
Faculty of Mechanical Engineering
Technical University of Sofia,
Specialist in Automation

1. INTRODUCTION

1.1. The subject of study: locked dynamical control systems

The subject of this work is locked automatic control systems (ACS) [1–18]. An example of such system can be any robotic arm assembly, as well as any part of robot manipulator. Such systems are needed not only in robotics but also in all areas of engineering, technology, industry, business, and science. Wherever there is mechanical control of a physical value of an object, there are such systems.

The main goal of this book is to develop methods of numerical optimization of ACS. Main tools are connected with software VisSim [19–20].

The software complex *Simulink* + *MATLAB* (along with the set of packages for their expansion *Toolbox* and *Blockset*), as *V. Dyakonov* rightly notes [49], proved too cumbersome for such relatively simple applications as modeling and optimization of regulators. This package takes up a noticeably large amount of memory, an excessively large library of blocks, most of which are unnecessarily specialized and are not required in most of the tasks to be solved; the created files also occupy a fairly large amount of memory, in total this amounts to several gigabytes. For this reason, there has recently been a sharp increase in interest in a small-scale but sufficient universal system of block imitation visual-oriented mathematical modeling, such as *VisSim*. This software has been created by the corporation *Visual Solution Inc*. (USA). The main developer of the software and the Head of the corporation is *Peter Darnell*. Along with the system itself, a number of packages for its expansion have been released, significantly increasing the already tangible capabilities of this system.

This “pearl” in the world of mathematical modeling programs has long attracted the interest of specialists in the field of mathematical modeling. For example, the corporation *MathSoft*, the creator of the famous and most popular computer mathematics *Mathcad*, not only ensured the docking of this system with the *VisSim*, but also began to supply *VisSim* with some versions of the *Mathcad* package [49]. The VisSim version can also be integrated with the *Simulink + MATLAB* system [49].

A major contribution to the distribution of the *VisSim* was made by the site www.vissim.nm.ru, created under the direction of *N. Klimachev* from the South Ural University, some versions of this program, for example, *VisSim* 3.0, are distributed free of charge, later versions can be used within 60 days for free, these versions can be obtained from the website www.vissim.com [69].

The objectives of the work are description and discussion of the developed modeling techniques, the example of research and optimization of such systems, development of recommendations for their further modifications and their successful application [21–42].

Modern production processes are inconceivable without ACS, providing stabilization of a number of important characteristics of these processes and their control by the prescribed rules of technology.

These systems are dynamic because in the calculation of such systems one must take into account the dynamic properties of all their elements and relationships. The action of these systems evolves over time and essentially depends on the dynamic characteristics of the system as a whole, which complexly depends on the dynamic characteristics of its constituent elements. When taking into account the feedback action it is inadmissible to neglect the dynamic errors in the signal path delays, dynamic dependencies, because such neglect will give an erroneous result.

We can illustrate this with the following two examples.

Example 1. Let us assume that a car is moving along the road and the driver follows the route with the help of the global satellite navigation system. The system indicates the correct direction of travel. The driver adjusts the car's route, so that the car is moving closer to the goal. A new possible route requires a new direction of movement. The entire system, consisting of a vehicle (object), the driver (regulator) and the navigation system (sensor of the output value) is closed. However, it is not necessary to consider this system as dynamic system, although its action takes place in time. The time it takes to update the position of the object and to adjust the direction of travel is negligible compared with the time that is required to achieve such a new situation in which the previous instructions in the direction of movement are no longer relevant. Delay in the adjusting of the course is so insignificant in comparison with the time of movement on the exchange rate that the system can become unstable as the result of this delay.

In addition, this system is not automatic, because a human is presented in the loop of its control. Systems with a human in the loop are automated but not automatic. This book does not consider such systems.

Example 2. Let us assume that a two-wheeled motor axis is an auto-balancing device. Such system is called balancing robot or Segway and includes position sensors devices, drives to the motor, the motor itself and the power supply of the entire system. The center of mass of this system is higher than the wheel axis of rotation. Therefore, in the absence of a control signal to the motor system capsize so that the center of mass located below the axis of rotation. If you want the center of mass to remain above the axis constantly, it is required to form the impact on the axis to this balancing robot to maintain a balance by moving forward or backward. If in such system the developer wrongly considers the time of receipt of the signal from the sensors, it will not be able to work steadily. The speed of all its components must be strictly coordinated with each other and match the pace of change random noise acting on the system. This system is dynamic. Such systems are the subject of this book.

1.2. Feature of ACS: negative feedback

Let us consider management of a structure **without feedback**, as shown in Fig. 1.1 and explain why **this structure is not applicable**.

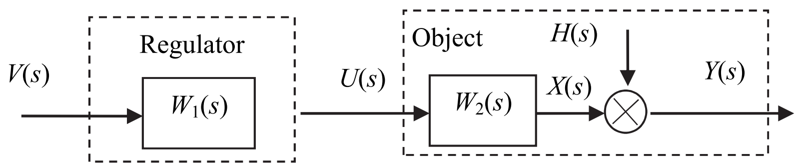


Fig. 1.1. Control without feedback

Any device or physical object can be treated as control object if it has output value and input for changing this value. The output value $Y(t)$ depends on the input controlling value $U(t)$. If there is not such a relationship, then the object cannot be controlled. If the relationship is strictly deterministic and the output depends only on the input signal, the control of such object is not among the tasks we have to solve, as it is fully described by its inputs as usual peripheral device, such as a stepper motor. The tasks of this book include control of objects, which operate in real environment and gets the impact of this environment. Along with the fact that the control signal can change the output value (it can also be called output signal), this value is also influenced by some external factors (disturbances). Therefore, the control task is to provide complete dependence of the output values on the prescribed value and the complete absence of its dependence on external factors (disturbances).

If the dependence of the output value from the input signal is linear, at least approximately, it is possible to describe it with the transfer function, which describes the ratio of the output value to this of the input signal. When describing this ratio, we use Laplace mathematical apparatus in order to take into account the dynamic properties of the object and other elements of the system. If the object model is non-linear, the transfer function is not suitable. In this simulation, this feature does not play a significant role, since software tools allow simulating both linear and non-linear elements.

Let the controlled object is linear and it can be described by a transfer function $W_2(s)$, which represents the ratio of the output signal $y(t)$ to the input signal $u(t)$ in Laplace transformation, i. e. $W_2(s) = Y(s)/U(s)$. It is necessary to provide that the output value become equal to the prescribed value. Let the prescribed value is given by the signal $v(t)$ and called "prescription". It is required to convert the prescribed signal into control signal in such way that the output signal would become equal to the prescribed signal, at least approximately. If the mathematical relationship between the output signal of the object and its input is known, then it is not difficult to calculate the desired signal, which would lead output value to the desired value. Since signals are best considered in the Laplace transformation where the output of a linear link is described as the product of the transfer function of the link to the transform of the input signal, it can be assumed that the necessary input signal in Laplace transform is the result of the dividing of the necessary output signal transform to the transfer function of the object. If the object would not be affected by any other influence except the control signal, it would be fair.

However, in reality, the output signal of an object is equal to the sum of the signal calculated in this way and the results of the action of all uncontrolled influences on it, which together can be described as some additive disturbance at the output of the object. In Fig. 1.1 this fact is displayed by the structure of the object model, namely: the action of the control signal is described by the transfer function of the object $W_2(s)$, the result of the action of the control signal is indicated by the virtual signal $X(s)$ corresponding to the signal in the form of the time function $X(s)$. In addition immeasurable disturbance $h(t)$ effect on the object. This signal is added to the virtual signal $x(t)$, so that at the output of the object we do not have the signal $x(t)$, therefore the outputs signal of the object is not signal $X(s)$, but the sum of this signal and the disturbance. Since the disturbance is unknown, the output value of the object is also unknown.

Example 3. We would like to control the heating of an aquarium, and we know that if we apply for a heater voltage of 10 V, then the temperature will increase by 1 °C. The transfer function of the object (excluding its dynamic properties) can be treated as gain of $K_O = 0{,}1$°C/V. If it is necessary that the aquarium temperature would be higher than the temperature in the room by five degrees, the prescription should be equal to $v(t) = +5°$ C. The regulator must convert this value into the appropriate voltage and the required conversion factor is the inversion of the according transfer function of the object: $K_R = 1/K_O = 10$ V/°C. In this case, the control signal is supplied to the heating element and will be equal to the product of the prescription to this coefficient, namely: $u(t) = K_R v(t) = 50$ V. That is, the regulator should apply to the heater the voltage of 50 V. At the same time, we do not take into account which temperature is in the room. In addition, other external or internal disturbances may act and that is increasing or lowering the temperature in the aquarium. It may also be that the mentioned ratio is not constant or dependent on the temperature or other conditions. Therefore it is impossible to know accurately even the temperature increment. Therefore, such system does not control the temperature, but only adds some effect with uncertain result.

Example 4. If in Example 3 we can measure the temperature in the aquarium and on the basis of this measurement change the voltage applied to the heater to provide the desired temperature at the outlet, then even ignorance about the transfer function of the heater would not prevent to solve the problem accurately. For example, we could heat the water up until it reaches the desired temperature, after which we would switch off the heater. Once the water temperature drops, we would again switch on, and so on. This operation can be made of automatically, not using a person by combining the temperature sensor directly with the switch. This principle is in the base of the action of the automatic electric heaters with the simplest elements of the automatic adjustment. Thus, even if heating is not as efficient as expected, the sensor will still not disconnect the heating power up until the temperature reaches the desired value.

Conclusion 1. In systems with feedback, the control may be sufficiently accurate even if the object model is known accurately enough, and even in case when this model changes over time due to some unknown factors.

Thus, only the negative feedback loop provides sufficient accuracy of control of the output values of the object. Fig. 1.2 shows this control loop in general. Automatic control theory is developed enough to control the linear stationary objects, but with the number of variables of the object and (or) with the growth of the order of the equations describing the relationship between the input and output signals, the analytical solution of this problem becomes extremely difficult. These problems arise when it is necessary to take into account nonlinear or delay element. In this case, it is much more successful to apply numerical methods based on simulation of such systems [20–42]. Simulation allows the selection, implementation and simulation test of the regulator, but without theoretically based and practically proven methods of these possibilities a developer cannot implement them effectively. Thus, it is relevant the design, development and justification of methods of numerical calculation of regulators for objects that contain high-order elements, nonlinear and transcendental elements. It is also necessary to test these methods to solve the regulator design problems for continuous technological processes with the objects characterized by the mentioned features.

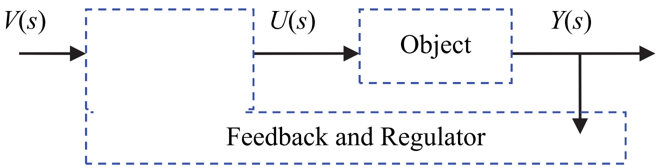


Fig. 1.2. Control with feedback loop

Design of the feedback system demands the development of a regulator and total structure of feedback. In addition, coefficients in these structures should be calculated or found with any other method.

The selection of the structure is often extremely simple. The feedback is used with a unite coefficient. It is put through a subtractor, and device, which usually has three separate control channel, is used as a regulator. These three channels include channel with proportional converting of the signal (proportional channel), channel with an integrating converting (integrating channel) and channel with derivative converting (derivative channel). Hence, the name of the regulator, containing all three channels is PID-regulator, of PID-controller, or PID, according the first letters of the three channels.

If the regulator does not contain any of these channels, the corresponding letter falls from its name, for example, remains the PI or PD. If some of the channels is presented twice, the letter can be doubled or given with an exponent, for example, PIDD or PID^2. If there is only proportional path, can be used historically established name "*K*-regulator", where "*K*" stands for easy gain, which, as a rule, should be large enough for the success of the tasks of the object control.

As a rule, control objects in industrial process lines are characterized by a significant delay, which cannot be neglected even in comparison with the phase delay of the minimum-phase part of the transfer function. In addition, non-linearity of the object produces strong influence to the object actions. Along with the presence of cross-linking, it requires consideration of objects like multiply and non-linear at the same time. Papers of many scientists are devoted to research of the methods of control of such objects [21–42].

In particular, control techniques are developed for multichannel objects using the method of the highest derivative of the state vector, the method of analysis and design of adaptive systems based on the principle of localization have been successfully used, methods of separation movements and some other methods and techniques are also developed and used [20]. Cases of linear objects are considered and studied deeply enough. As a rule, the decision of the task of the regulator design in this case is solved analytically, mostly in consideration of the space object states. For objects that contain delay and (or) non-linearity, it is not always possible to use analytical methods due to significant increase in the complexity of the problem. Problems of control of higher order multivariable objects whose matrix transfer function has a greater dimension and thus contains elements of higher order links and objects that contain nonlinear transcendental or links, are extremely complex. Most often, they cannot be solved analytically or their resolving with the use of the known methods is extremely difficult, and it requires enormous computing resources or a lot of time, which makes these tasks insoluble by known techniques. The use of different simplifications of the problem leads to the fact that the resulting solution is far from the desired, or the difference of practical implementations from the theoretical results is unacceptably high.

Numerical simulation and numerical optimization are increasingly used due to the development of mathematical methods and special programs, as well as due to increased computing performance. These methods allow finding acceptable solutions to the simplest examples of this class of problems, although a universal method of solving these problems do not exist. Application of the most effective programs for modeling and numerical optimization allow making significant progress towards addressing the growing challenges in this field. However, for this purpose methods of applying these programs to these problems are necessary: the lack of practical methods and theoretical justification for them unreasonably curb the proliferation of numerical methods for optimization of ACS.

In this regard, development of theoretical basis of methods and practical algorithms for the design of regulators is relevant. It should provide high-quality control one- and multi-dimensional objects that contain delay elements, nonlinearity and other characteristics listed above, which not allow the analytical calculation of the regulators.

Closed dynamic automatic control systems are necessary in almost all fields of science, technology and industry. They provide stabilization or change in the required law the most important characteristics that are necessary for technological or research operations and allow control transport, robots, chemical, biochemical and physical processes.

Achieving the required accuracy of these values is provided only with control loop with negative feedback [21–42].

1.3. The task of regulator calculation

Any ACS contains the following parts: a) sensor, measuring the output value; b) regulator (controller); c) actuator (the device acting to object). The regulator is connected between the sensor and the actuator; it converts the sensor signals into the control signals based on the prescription. With all this, the regulator is designed to ensure the stability of the closed system and to provide the required static and dynamic accuracy. The dynamic properties of the regulator are set entirely by the developer, in contrast to the dynamic properties of the other elements of the system, which are determined by the conditions for their implementation and cannot be arbitrarily changed. The regulator is intended to supplement or correct the dynamic and static properties of other elements to produce high quality to the whole system. Design and calculation of control is the problems of high complexity because it requires knowledge of the object model and all other elements of the system, good knowledge of control theory and special chapters of applied mathematics. Analytical calculation of regulatory is time-consuming, even in the case of single-channel linear objects. For more complex objects it is sometimes just not feasible. Therefore, approximate methods are used, numerical methods or other simplified procedure [21–42]. The development of computers and software for the simulation and optimization has given the opportunity for more successful and easy solution to these problems.

This book is dedicated to the methods of computer-aided design of regulators using *VisSim* 5.0 software package and its more recent modifications [21–42].

The principle of the stabilizing effect of negative feedback with a high coefficient is widely known [19–20]. **Feedback is negative** when occurring in the loop disturbance of the stabilized value (a deviation from the equilibrium state) gives rise to signals within the loop, which at its point of generation act in the direction opposite to the action of the disturbance. For example, a decrease in temperature of the object in the thermal stabilization system, regardless of the disturbing external factors should produce such action of the regulator, which will lead to an increase in temperature of the object. Due to the properties of the feedback this increase will be exactly as what is necessary to maintain the temperature of the way it should be, as if this disturbance was not exists. If, on the contrary, due to external causes of the temperature rise, the feedback compensating action causes cooling and the temperature will result in such as it is prescribed by the prescribing signal.

However, the implementation only of the principle of negative feedback is not enough for the effective operation of the regulator. At certain ratios of speed and gain of the system components **instability of the system** can occur. It lies in the fact that at some frequencies even the smallest deviation from the equilibrium state generates an action that increases the deviation; i. e. the system with negative feedback behaves as a system with positive feedback. Due to the delay in the spread of the signal in the control loop and object in a certain frequency range negative feedback can indeed be positive, since the phase shift at 180º in transmission of harmonic signal is equivalent to a change of the sign of the signal. Positive feedback even in a limited frequency range will cause a loss of stability of the entire system, if the coefficient of the connection is greater of unit. For a small gain (less than one) frequency, local (in a narrow frequency band) positive feedback will not disturb the stability of the system.

Thus, the system includes a serially connected sensor, regulator, actuator and the control object. It should have a well-defined frequency response, to provide not only the desired gain (in charge of the depth of the disturbance suppression), but also the stability and the necessary stability margin.

Since the properties of object are set, and the system designer cannot change them, the regulator must be designed based on these properties to ensure required stability, accuracy and speed.

Therefore, it is necessary to know the mathematical model of the object and methods of calculation of regulators on this basis. For the mathematical model of the object, as a rule, it is used the researching procedure consisting of the formation of test actions, the study of object responses to these impacts and finding mathematical model of the object from these data. These issues are handled by one of the technical sciences, called the **identification of objects of control**.

If a mathematical model of the object is known, on this basis, taking into account the requirements of the system regulator should be calculated. These questions are decided by **automatic control theory**, which is sufficiently well developed for the tasks of linear stationary object control. However, with the growth of the order of the differential equations describing the relationship between the individual signals, the analytical solution of this problem becomes extremely difficult, even with the development of modern computing methods and tools. These problems are compounded when it is necessary to take into account nonlinear elements or pure delay units. In this case, it appears to be more successful application of numerical methods are based on mathematical modeling (simulation) of such systems. Mathematical modeling allows the selection of the regulators and the simulation of the system with this regulator (calculation of all its signals). In addition, modeling allows changing the settings to the regulator if the processes in the system are not satisfactory. However, without theoretically based and practically proven methods these features cannot be implemented effectively.

Implemented design, development and study of methods and techniques of numerical calculation of regulators for most practical objects that contain high-order elements, nonlinear and transcendental elements, makes effective use of the program VisSim [32–34] for the calculation of regulators to ensure the required stability and accuracy [35–42]. Main statements and results of this procedure are set out below.

Transfer functions are used to describe linear inertia objects (i.e., their output signals are depending not only on the input ones, but on temporal characteristics also), whereas for non-linear objects description such structures as transfer functions should be complemented by appropriate nonlinear characteristics that do not have their own inertia properties. By this method it is possible to describe and simulate nonlinear objects, but it is very difficult apply or even impossible to use analytical description for this task because the object model comprises units described in various forms of mathematical models. In particular, the best model for linear object is Laplace transform in the operator domain, or it can be described in the form of differential equations. For non-linear object, the best description form is in the field of real-time signals, time varying. The Laplace transform is description of the signals not in the time domain but in frequency domain.

1.4. Questions for self-control

1. Draw typical structure of the system without feedback and with feedback; explain the difference between them from the point of view of the mathematical models and their using.
2. Give the reason, why in systems with feedback, the control may be sufficiently accurate even if the object model is known accurately enough, and even in case when this model changes over time due to some unknown factors.
3. Explain the terminology of *PID*-regulator, as well as *PI*-, *I*-, *PD*-, *K*-regulator. How do you thing, what can be *ID*-regulator? Has this name some practical sense?
4. What is the difference between *P*-regulator and *K*-regulator? Explain the meaning of the names PIDD, PID^2, PIID, PI^2D PIIDD, PI^2D^2-regulator?
5. What do you think can mean the name $PI^{1/2}D^{1/2}$-regulator?
6. What means the negative feedback and positive feedback? Can neutral feedback exist (without polarity)?
7. What means stability or instability of the system? Give illustrative examples.
8. What is identification of the object? How you treat parametric identification, structural identification, full identification, restricted-frequency identification?
9. Which are the useful features of the stable system? What about the instable one – can it be useful?
10. How do you think, the stable oscillation (for example, in some in generator) can be realized as a kind of instable negative feedback?
11. What is linear system? Can you give examples of linear and nonlinear ones?
12. Can be Laplace transform used for linear object (system)? What about non-linear one?

13. If the system has deep negative feedback with small phase delay (less than 180°) but in some local frequency range (not in the highest frequencies), it has local virtually-positive feedback due to the concrete phase delay, is this system stable as a whole?

14. How do you think, why the software VisSim is chosen as the base for the decision of the tasks of the optimization of the feedback control?

15. What is numerical optimization of the regulator? Can be other kinds of the optimization? How you explain structural optimization?

16. How do you think, is the term "more optimal" (or "less optimal") correct? Did you meet such terms in civil conversations? Did you meet such terms in science publication?

17. If the task of the optimization would change in its formal indexes, whether then the result of the optimization will change? If yes, then the optimization task has several variants of the decisions, isn't it? With these aspects, revise your answer to the previous question. Do you insist on it? In any case, explain your opinion.

If you have difficulties in answering some questions, return to them after studying the following chapters. Other questions are given in the end of the book and they are mixed on the order. It can be used as questions at exams.

2. TERMINOLOGY AND MATHEMATICAL TOOLS OF TAC

2.1. Basic requirements and mathematical tools for the system

The dependence of the requirements of the regulator on the requirements for the system is big. Different areas of control theory are dedicated to the study of this relationship for the different classes of objects.

The most indicative characteristic of closed linear dynamic systems is the frequency amplitude response of conditionally open loop that includes all elements of the system, but for the sake of completeness, phase response of this circuit is also required. Conventionally, all the elements, which frequency properties cannot be changed arbitrarily (sensor, actuator etc.), refer to the object that allows to consider the system as connected in series in the loop regulator and the object only.

For linear systems, tool of complex frequency characteristics or the associated apparatus of Laplace transforms are the most effective. The argument of the Laplace transform is the complex parameter s, and for the frequency characteristics frequency is used, which is its imaginary part. Fig. 2.1 shows typical structure of a closed loop control system, where the frequency characteristics of individual links are written in rectangles, and type designation of the signals are used in the system.

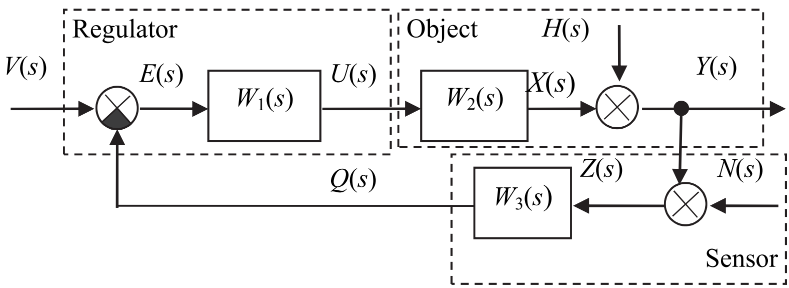


Fig. 2.1. Structure for calculation of ACS

Namely: $y(t)$ is the output value of the object; $v(t)$ is the prescribed value of the output value; $u(t)$ is the control signal supplied to the control object;

$h(t)$ is disturbance acting on the object, which value is converted to units of output value; $x(t)$ is the state of the object, i.e. This is a virtual output value, which would be in the absence of disturbance. In addition here $e(t)$ is the control error; $n(t)$ is the noise of the measurement of the output value; $z(t)$ is the result of the measurement of the output value; $q(t)$ is value of the sensor of the output signal $y(t)$. In addition, traditional replacing of lower case letters to upper case is used together with replacing of functions of time on their conversion operators, for example, $V(s)$ is the transformation of the signal $v(t)$. Note that the Laplace transform of the unit constant is $1/s$. Since, as a rule, the Laplace transform values of the input, output, and intermediate signals are not used for calculations, but only their relationship is used (i. e. transfer functions), therefore the practice of use of the modified transform has been established. Namely the Laplace–Carson, which is obtained by multiplying of Laplace transform by s, so that the image of the constants is also a constant. It does not effect on the values of transfer functions, so we will use the terminology of Laplace transforms.

The structure in Fig. 2.1 is graphical representation of the relationship of signals, which can be replaced by the following equations:

$$E(s) = V(s) - Q(s); \quad U(s) = W_1(s)E(s); \quad X(s) = W_2(s)U(s);$$

$$Y(s) = X(s) + H(s); \quad Z(s) = Y(s) + N(s); \quad Q(s) = W_3(s)Z(s).$$

To solve this system with respect to any of the values within the loop (for example, the dependence of Y from V, N and H) as a result inevitably appears a rational fraction of the transfer functions (readers are invited to draw conclusions on their own).

If the denominator of this fraction $1 + W_1(s)W_2(s)W_3(s)$ become zero, the whole fraction has an infinite value. This means that any arbitrarily small input signals produces “arbitrarily large” output signals, and taking into account possibility of the physical realization, this means that the system instead of being in steady state moves to the maximum possible deviation from it or oscillates around it with the maximum achievable amplitude.

To avoid this situation, the regulator is necessary, calculation of which this book is devoted.

To investigate the stability of the system it is necessary to examine the relationship of dynamic models of the elements included in the system. If identical mathematical models describe two systems and the output signals should are identical then the input signals are also identical. *VisSim* software allows simulation of the most known dynamic links, as well as implements the connections of signals from the outputs to the according inputs with the goal to generate the necessary input signals. Therefore, simulation run produces graphs of transient process identical to real signals in a real system, subject to full compliance with all mathematical models of their prototypes, i.e. the elements of the real system.

The most common elements of the model are linear dynamic links and are described by rational functions of *s*.

2.2. Requirements for the possibility of the physical realization of the model

All elements of the system model must be physically realizable. This condition formally requires that the degree of the numerator is less than that of the denominator. Indeed, it should be noted that the argument s is some analogue of frequency. With increasing of frequency, the transfer function of individual elements is not necessarily falls, in observation of it far enough along the frequencies axes, the transfer function of any real object is sure drop lower than arbitrarily low values, i.e. becomes almost zero. For any object it is always possible to specify such big frequencies that the transfer function on them is not only small but practically zero. It means that the response of the object on these frequencies is much less noises of the measuring.

Frequency responses are usually represented in a logarithmic scale, so the chart achieves "virtually infinite" quickly, i.e. the plot simply presents very large frequencies and the transfer function values. It also simply presents practical zero, i. e. very small values of these variables. Hence, values of frequencies, many times exceeding the frequency border of the object transmission are also clearly visible on the logarithmic graphs. Thus, the right side of the chart of logarithmic frequency response (LFR) of any real link should fall down: the more right, the more down. This is achieved only under the condition that the order of the denominator of said transfer function is greater than the order of the numerator.

Sometimes, designers neglect this requirement because the bandwidth of these links may be significantly wider of the transmission bandwidth of the object. Separate elements damping can be neglected only in comparison with damping of other units included in the same loop. If any link is the only link in the loop under consideration, therefore there is no other inertial element in this loop, in comparison with which damping of the element could be negligible. In this case, it is unacceptable to describe said link as link without inertial properties.

Even if designer consider that the difference of the order of the denominator from this of the numerator by only one or two orders, it would be also not enough adequate model for stability analysis of almost any real loop with negative feedback.

However, if there is delay element in the loop and if it is known that, the effect of the following time constant of the model phase-frequency response (PFR) is substantially less than the delay, then it is permitted to restrict consideration of the dynamic elements of the object by the first or second order only.

Example 5. Let us try to imagine the mostly high-speed object. Assume high frequency electronic signal amplifier based on one stage on microwave transistor. For example, high-frequency transistors can be characterized by the gain bandwidth of up to 1 GHz, i. e. 109 Hz. Assume that the gain of such a transistor is equal to 2, which means that the transfer function of the transistor in the first approximation has the following form:

$$W_T(s) = \frac{2}{Ts+1} . \qquad (2.1)$$

Let $T = 10^{-9}$ *s*. Inside the bandwidth of the transfer function the amplifier gain is close to two. Let substitute the value of the angular frequency $\omega = 10^{12}$. In this case, the transfer function is approximately 0.002. This means that in said frequency the amplifier has the said gain. However, said frequency is the frequency of optical radiation. Electronic devices, except optical devices do not transmit optical signals. That is actually the specified amplifier cannot transmit any signal, hence its transmission coefficient must be exactly zero. No final-order model will not allow such a result, according to which the transfer coefficient of the object at this frequency is exactly equal to zero.

Example 6. If Example 5 seems not conclusive, it is possible to consider more high frequency, for example, $\omega = 10^{15}$, $\omega = 10^{20}$ and so forth. There can be no doubt that for all practical example we can find a frequency at which the transmission coefficient should be not just a small but exactly zero, that is, the response of the object at this frequency cannot exist. No linear model of finite order filter gives such a property of the object model.

Conclusion 2. The description of the object in the form of a linear finite-order filter is always an approximation of the real model.

Conclusion 3. Any model in the form of the transfer function of finite order is an idealization. With increasing in the frequency, the relative error of this idealization (the ratio of the error to the actual value of transfer function) is growing exponentially or faster. The real meaning of the transfer function tends to zero faster than the corresponding value of such model.

2.3. The choice of the regulator structure

The choice of the regulator structure depends on the object model and on the resolving task. The simplest structure is obtained if in the structure of Fig. 2.1 the transfer function of the sensor together with a device converting it signals present, as a link with a single gain, and if the term “regulator” is narrowed down to the concept of the transfer functions of its sequent part, as shown in Fig. 2.2.

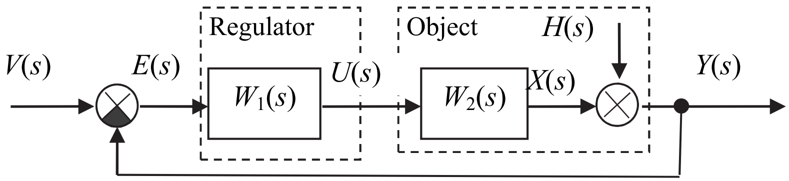


Fig. 2.2. The Simplified Structure of ACS

The most complex and generalized scheme will occur, if on the contrary, the term “regulator” would be extended to include the signals from all possible points in the system, as Fig. 2.3 shows.

The complex structure of Fig. 2.3 may have an advantage in some special cases. However, it is not free from several drawbacks. Firstly, the disturbance signal is usually not available for the measurements, but if it is available under the terms of the problem, in practice it is performed only partially. Secondly, most of the main problems is to control precisely disturbance suppression, however without loss of generality it can be put $v(t) = 0$, $V(\mathrm{s}) = 0$.

It follows that the input element number 3 is zero, so that the element can be withdrawn from the structure, as well as the element number 5 due to inaccessibility of the signal $h(t)$. Element number 4 is included in the same way as the element number 1, but without inversion, since the first input of the subtractor device receives nothing, it can be replaced by the inverter, or completely removed by changing the sign of the transfer function of element number 1. In this case, the structure of Fig. 2.3 is completely equivalent to the structure of Fig. 2.2.

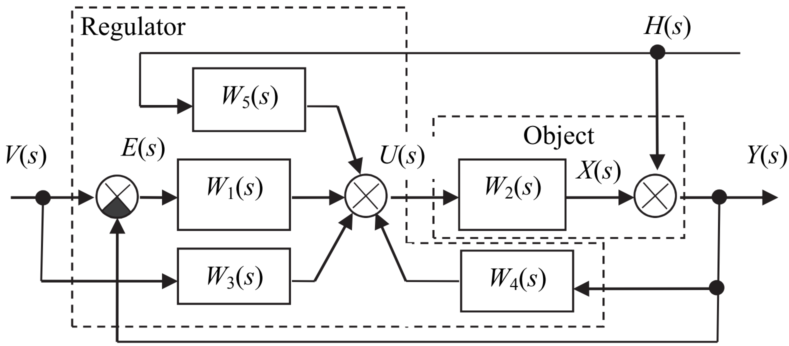


Fig. 2.3. Common Structure of ACS

Typically, the output value is not directly available for measurement, but it only can be estimated using a sensor. Sensor action should be also described with the transfer function of more complex model. It is advisable directly between the sensor element and the subtractor to use the converter, which compensates for the transmission coefficient of the sensor so that the subtractor would use the calculated output value of the object. Series connection of the transfer function of the sensor and the compensator is approximately equal to unity at low frequencies. Therefore, such a system can only be treated as system with unit feedback. Of course, in this case the dynamic properties of the sensor are missed from consideration. Wherever possible, the properties of the sensor must be considered in the structural scheme of the system, and where they are not considered, it is assumed that they are negligible in comparison with the inertia and the object. If this assumption is wrong, then the regulator calculation will be made mistakenly.

2.4. Regulator with fixed coefficients

Design of ACS, as a rule, takes place in four stages: 1) definition of the object model; 2) design of the regulator structure; 3) calculation of regulator parameters; 4) testing of the system and adjusting of the regulator parameters if necessary.

Often, instead of designing of the structure the most typical one is used, sometimes even without sufficient justification. For several reasons, the most commonly used structure among regulators with constant coefficients is a PID regulator, i. e. regulator comprising proportional, integrating and derivation control channels. Such a regulator can be set with the following transfer function:

$$W_P(s) = K_P + \frac{K_I}{s} + K_D s\,. \tag{2.2}$$

Here, K_P, K_I, K_D are constant coefficients which are required to determine with the procedures as a result of numerical optimization. Equation (2.2) can be written in different forms, including a reduction to a common denominator. If any of the channels is not used, only the letters remain in the title, which correspond to the used channels, for example, PI or PD regulator.

With proportional regulator, the shape of the logarithmic frequency response (LFR) of the open loop coincides with this of the object. The regulator only moves it up and down parallel to itself. If the ratio is greater than one, the graph is moved upward by an amount corresponding to $20\lg K_P$, if less than one, then the graph is moved downward.

The integrating channel increases the slope of the whole LFR (if only it is included to the regulator) or only of its low-frequency part without changing the type of midrange and high frequency parts of it (if there are also proportional channel in the regulator). Increased slope is asymptotically equal $\Delta\Theta = -20$ *dB/decade*. If the slope was zero, it will be equal to this value, if it was negative, it will increase by this amount.

The derivative channel reduces the slope of the high-frequency part of LFR, namely it adds to it a positive value of 20 *dB/decade*. If it is negative, then it is reduced in magnitude to that value.

Integrating channel reduces static error to zero. Derivative channel in general provides the best margin of system stability and reduces overshoot. There is not strict relationship between the stability magnitude (and the system quality) from the coefficients in equation (2.2), therefore numerical optimization method of calculating these coefficients or other method is necessary. Although the regulator (2.2) is most commonly used, and it is often the most effective, in some cases, such regulator is not applicable, and in some cases, one of its channels hinders, rather than helps. The advantage of the numerical optimization procedure is also the property that if the regulator contains an excess channel, then the correct procedure performed with correctly selected quality criteria should lead to the fact that the coefficient of hit channel will be negligible. For example, if the integrator channel is not required, the obtained coefficient will be extremely small.

Another form of the regulator (2.2) is:

$$W_P(s) = \frac{K_П s + K_И + K_Д s^2}{s}. \tag{2.3}$$

More sophisticated regulators may include additional derivative and (or) integration. For example, if the regulator contains both these additional channel, its transfer function might look like this:

$$W_P(s) = K_П + \frac{K_И}{s} + \frac{K_{И2}}{s^2} + K_Д s + K_{Д2} s^2. \tag{2.4}$$

Such regulator is named PI2D2 or PIIDD-regulator.

2.5. Regulators with a non-integer integration and differentiation

It is also possible to record a transfer function in which the orders of the exponents of the argument s in the numerator and denominator are positive non-integer but fractional values. For example, the following form of the regulator transfer function can be found in literature [21]:

$$W_P(s) = K_П + \frac{K_И}{s^{\mu}} + K_Д s^{\lambda}. \tag{2.5}$$

Here μ and λ are positive values; at least one of them is fractional, or both of them can be fractional.

In the literature, there are the following reasons for this decision.

1. Fractional integration allows increasing of the slope of the high-frequency part of LFC on absolute value, e. i. more than 20 *dB/decade*, but less than 40 *dB/decade*. The greater the slope, the faster LFC increases with moving towards lower frequencies of LFC. It means more effectively suppressing of the noises in the middle frequency range. Dual slope of 40 *dB/decade* is inadmissible, as in this case, the system becomes unstable, and a single slope of 20 *dB/decade* it may be not sufficient.

2. Regulators with arbitrary values of μ and λ are seen more universal, since acquiring two additional parameters. Therefore, it is assumed that the possibility of such regulators is higher.

However, the transfer function (2.5) is easy to write but difficult to implement in a real regulator. If it is implemented with the help of elementary units of derivation and integration, it should be noted that the very these units in digital form are implemented using addition and subtraction operations. In particular, let us consider the simplest "half" of integration:

$$W_P(s) = \frac{K_И}{1 + s^{0,5}} . \tag{2.6}$$

For its implementation, it is necessary to carry out an approximate realization of the transfer function in the following form:

$$W_P(s) = \frac{K_И (1 + \tau_1 s)(1 + \tau_2 s)...(1 + \tau_m s)}{(1 + T_1 s)(1 + T_2 s)...(1 + T_n s)} . \tag{2.7}$$

Any operation or fractional derivation or integration may be implemented using only operations of integer integration and differentiation, and the approximation is satisfactory only to a limited bandwidth. Even very simple recorded regulator is implemented using a very large number of mathematical operations.

A reason for refusal of fractional integration are that twice or greater slope in the mid-field, passing in a single slope at high frequencies, is much more effective.

A convincing argument against regulators with fractional integration and (or) derivation is based on the fact that much more simple structure of regulator, for example, the form (2.4), can provide much better results than with this complex structure as (2.7) or more complex, which approximates not fractional integration or derivation.

2.6. Classification of developing direction of methods of the designing of ACS

The modern theory of regulators' design for multidimensional continuous technological objects containing in their model nonlinear and delay links is developing in several important directions. Among the methods of solving these problems, the following ones can be specified:

1. **Analytical** (mathematical) **methods** of research of objects and systems and of design of regulators.

2. **Tabular methods** of tuning.

3. **Empirical tuning** of real regulator.

4. **Mathematical modeling** for the study of properties of the object and system with the regulator, which has been found by and method.

5. **Numerical methods for optimization** of the locked loop model or a real system with the object based on built-in routines of modern software packages for modeling and mathematical calculations, including numerical methods, implemented by special devices designed as an “intelligent controllers” or “smart regulators”.

Analytical methods for tasks of control of multidimensional objects containing significant nonlinearity, delay and higher-order elements, as a rule, are not sufficient.

Tabular method (Ziegler-Nicholson method, Cohen-Kuhn, and similar) assume on the base of found characteristics of the transient process with the help of some empirically derived relationships calculation of the regulator coefficients. These methods are based, as a rule, on the assumption that the object model is the filter of the first or second order with an element of pure delay. Since this hypothesis is untenable in the most cases, these tabular methods work only in exceptional cases, but even in these cases they are far from optimal. In some cases tabular methods lead to the production of unstable systems, in other cases they give excessively large overshoot; cases of receiving systems with good quality using them are rare, but these methods are cited in the literature are still widely. Many authors mistakenly refer to these methods as the best ones.

The **empirical tuning** is applied in practice often enough, but at the same time received results are far from optimal, which requires the development of a more reasonable methods.

Numerical optimization is the most effective method of regulators tuning, if the object model is stationary and known with sufficient accuracy. It uses a mathematical modeling (simulation) of the system that contains the object and the regulator.

There are also papers with references to some other methods, which essence is not disclosed. This happens for one of two reasons: either the implemented algorithm in specialized devices is “know-how” of the developer, and therefore it is not published, or the authors of the papers do not know these algorithms, as they themselves are only the users of purchased regulators or systems. Such algorithms can be compared with other algorithms only on the achieved results. However, it not sure when it is based on the validity of theoretical assumptions, methods and techniques. Papers describing only the results without disclosing methods for their preparation do not represent scientific value.

Therefore, the most promising is the development of numerical methods for optimization of the closed loop model or a real object based on built-in routines of modern software packages for simulation and mathematical calculations. This development requires a theoretical justification of methods and techniques and the development of algorithms and tools for their implementation.

The starting premise for the development of such methods is the analysis of the results of the research systems obtained with regulators synthesized by known analytical methods for simplified examples.

2.7. Adaptive and self-tuning regulators

Along with the structure that uses the constant coefficients, the structure with variable coefficients can be used s well. Special devices can control changes of these coefficients, or they may be changed under the influence of chronometric devices. It allows creating of **adaptive regulators** and other types of complex regulators, such as **self-tuning regulators**.

Adaptive regulator should change their settings so that the system retains its original good quality when the object model is changing.

Self-tuning regulators should implement their own system configuration (automatically) **without the initial setup**. Self-tuning regulators can perform the function of the adaptive controller. Not every adaptive regulator can function as a self-tuning one. The adaptive regulator can be pre-configured to provide the desired quality of the system, but the object model may vary during its operation.

Adaptive regulator should contain a means of determining of the model of the object during the operation of the system. Determination of model can be carried out by analyzing of the relationships between changes in the input and in output signals. During the functioning of the system, it is often required that the object output signal would be constant. In this case, the problem arises that if the output does not change, the model of the object is difficult for definition, and if it is changed, the system is not functioning as desired. This complicates the design of adaptive self-tuning regulators.

2.8. Adaptive regulators and prospects of this approach

Adaptive regulators should change its parameters to provide the required quality of the system in conditions when the parameters of the object are changing. One of the essential features of such systems is to analyze any formal indicators of stability and quality of the system for decision-making on changes of the regulators settings. This approach allows compensating changes in objects' parameters within a certain range, usually known in advance. Thus, the design of an adaptive systems makes to exclude the last stage (test and correction) at the expense of more complex operating algorithm. It also provides the system with new properties, which are less dependent quality and accuracy of the system on uncontrolled change of object parameters, if the rate of change of parameters is considerably less than the rate of the transient processes.

The advancing of this approach can offer the development of such control algorithms that do not require determination of the entire object model completely or at least allow for mistake in identification of the object model compared to other methods. This is justified by the fact that the identification procedure itself is very laborious and costly. Therefore, the automatic regulator tuning can be less time consuming and more efficient than the sequential execution of the above four stages of experimental identification and analytical design.

Example 6. Let us consider two object models, which LFR differ only in the low-frequency field:

$$W_1(s) = \frac{1}{(1+200s)(1+0,4s+0,1s^2)}. \tag{2.8}$$

$$W_2(s) = \frac{1}{(1+200s)(1+0,4s+0,1s^2)} + \frac{0,001}{s}. \tag{2.9}$$

The responses of these objects on a unit jump step differ, as can be seen from the graphs in Fig. 2.4. Since the high-frequency part of these objects is the same, these objects behave almost identically inside the structure of automatic control system. Indeed, the structure in accordance with the simulation of Fig. 2.5 gives practically identical transient processes as Fig. 2.6

shows. In this structure, both object variants are simulated, depending on the presence or absence of connection of the lower branch (integrator) to a summing device.

Conclusion 3. The low-frequency part of the frequency characteristics of the object cannot have an appreciable impact on the transient process in the system, that is, on the stability of the system.

Conclusion 4. For successful operation of the adaptive loop is enough to know the high frequency and middle-frequency response of the object.

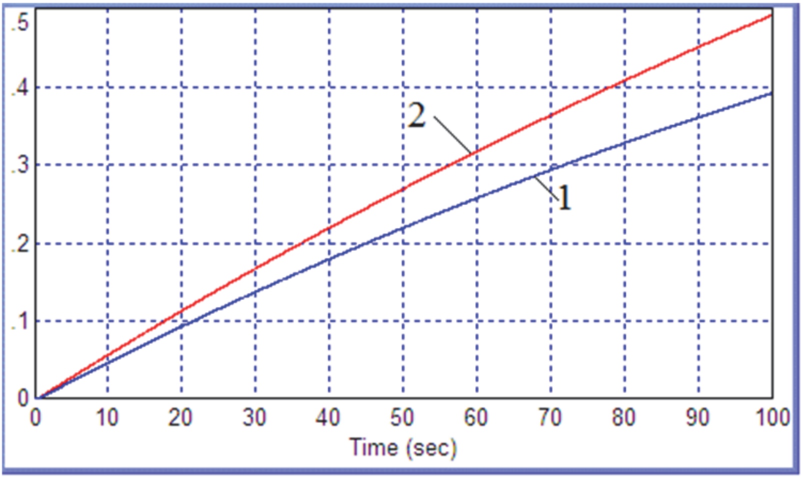


Fig. 2.4. The response to the unit step of the blocks with the transfer function: (2.8) - upper line 2; (2.9) - lower line 1

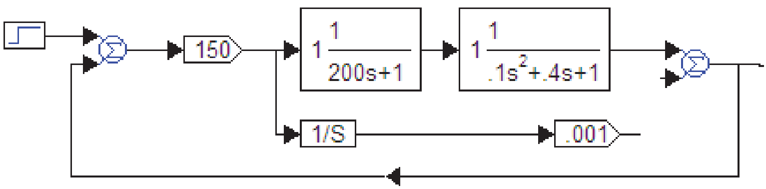


Fig. 2.5. The structure of the system for simulation of the object according equations (2.8) and (2.9)

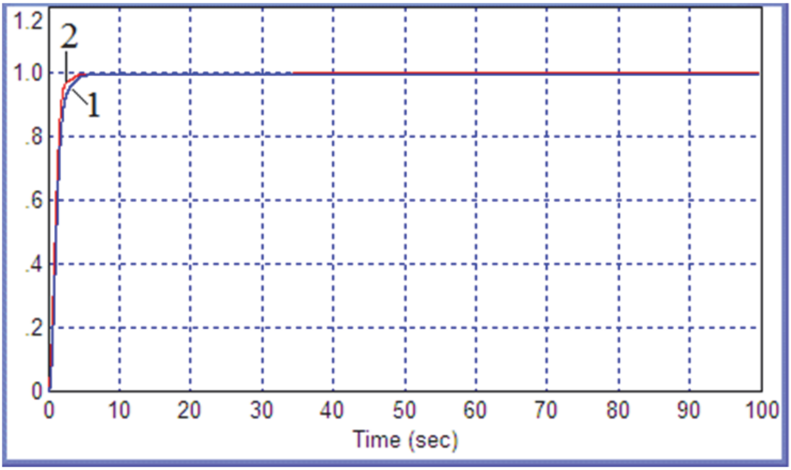


Fig. 2.6. Response of the system for Fig. 2.5 as response at the unit step change of the prescription: curve 1 is with the object (2.8); curve 2 is with the object (2.9)

The formulation of the problem of designing an adaptive system

That, what is claimed

1. The known model of the object or it can be determined as a result of the test cycle. It is also known that in this model, at least one of the coefficients vary in a certain limited range, the rate of change is limited.

2. There are requirements for a locked system.

That, what is required

It is required to design the regulator, such that the system meets the requirements of claim 2 under conditions of claim 1.

Solving methods

The regulator parameters can be varied during operation of the system under the action of a device, such as a signal analyzer system. Also, the regulator parameters can be changed under the influence of the timer. The structure of the regulator may also change, i. e. its mathematical model can be tuned or switched. In addition, preliminary tuning or calculation of the regulator can be used.

2.9. Self-tuning regulators

Empirical tuning of regulators are usually carried out without preliminary identification of the object model, only on the basis of some a priori knowledge of the object and certain statements of the theory of automatic control. The formalization of this process allows resolving the problem of the feasibility of adaptive systems, i. e. systems performing regulator implementation without identifying the object model, but only based on the form of the transient process or even on a part of it. If the regulator tuning process does not even require initial identification for the calculation of the regulator, the development and justification of such methods and design techniques allows developing of a new direction of adaptive systems, namely, the design of self-tuning systems. In the view of the capabilities of modern digital systems, it allows significantly reducing of the complexity of the design of the individual regulators by reducing the workload of the designer. This is achieved by increasing the complexity of algorithms of action of regulators that according this criterion can be considered as intelligent systems. Therefore, the greatest prospects of self-tuning algorithms are connected with the spread of digital regulators.

Self-tuning systems are often equated with adaptive ones, however, it is possible to specify some of the differences between these system types. In the first case, the main functional feature is tuning that is carried out, including when the designer starts the system first time. In this case it is carried out continuously during the operation of the system, but preliminary first adjustment is not included in this task. If the system solves both of these problems, it should apparently be attributed to the self-tuning system.

The urgency of this direction is determined by the need to development of new systems, the lack of highly qualified specialists for this, the difficulty of identifying of large class of objects, much more than an iterative tuning of regulators and a substantial simplification of the identification of the object in the case of existence of stable system achieved with iterative method.

In the future, performance of the regulator tuning before the identification of the object due to the iterative procedure simplifies the task of calculating it as a whole and reduces the cost of development of regulators.

To use these potential benefits to develop the theoretical foundations are suitable, as well as methods and algorithms of resolving of this problem, which would lead to the immediate solution of the tasks performed by software rather than by highly qualified operator.

The formulation of the problem of designing self-tuning system

That, what is claimed

1. Object model is unknown, it may be determined as a result of the test cycle, since at least one of the coefficients vary in a certain limited range, the rate of change is limited. In addition, the model may be stationary but not known.

2. There are requirements for a locked system.

That, what is required

It is necessary to design the regulator, such that the system meets the requirements of claim 2 under conditions of claim 1.

Solving methods

The regulator parameters should be changed from the start and during operation of the system under the action of a device, such as a signal analyzer system. It may also change the structure of the regulator, i. e. its mathematical model. Preliminary tuning of the system is not intended.

2.10. Robust regulators

If the object parameters are not accurately known, or change are uncontrolled but remain in a certain range, the task of designing such regulator can be supplied, which will provide acceptable control of the object for all combinations of object parameters that are within acceptable values.

This problem may be seemed insoluble. Even if the regulator has been found, ensuring of stable control with the extreme values of the object parameters, it may be that a combination of unfavorable parameters of object with the specified regulator violate the stability of the system.

In general, terms “favorable” or “unfavorable” for the combination of process parameters has a right to exist, if we mean that favorable combination allows provide more precision and more speed system. Unfavorable combination then call the combination that makes sacrifice performance and (or) the accuracy of the system for the sake of stability. If robust regulator is designed for the system, the favorable combination of parameters can be just as likely to break the stability, as well as unfavorable because “favorable” combination of parameters in this case should be related not to the potentially achievable speed and accuracy, but with the quality of the system for a given fixed regulator.

Robust controls can never be optimal, as well as the optimal regulators (for fixed object) may not be robust.

The method of numerical optimization of regulators for the ensemble of objects with a group quality criterion can allow calculating the robust regulator, which will be discussed below.

Example 7. Consider the object model:

$$W_1(s) = \frac{\exp(-\tau s)}{(1+47s)(1+0,5s+1s^2)} . \tag{2.10}$$

Suppose it is known that the delay unit time constant τ varies from 50 s to 60 s.

Let the system uses a PID regulator (2.2) with the following coefficients: $K_P = 0.562$; $K_i = 0.0087$; $K_D = 2.78$. Fig. 2.7 shows the transient processes for different values of the time constant: $\tau_1 = 50$ s, $\tau_2 = 55$ s, $\tau_3 = 60$ s.

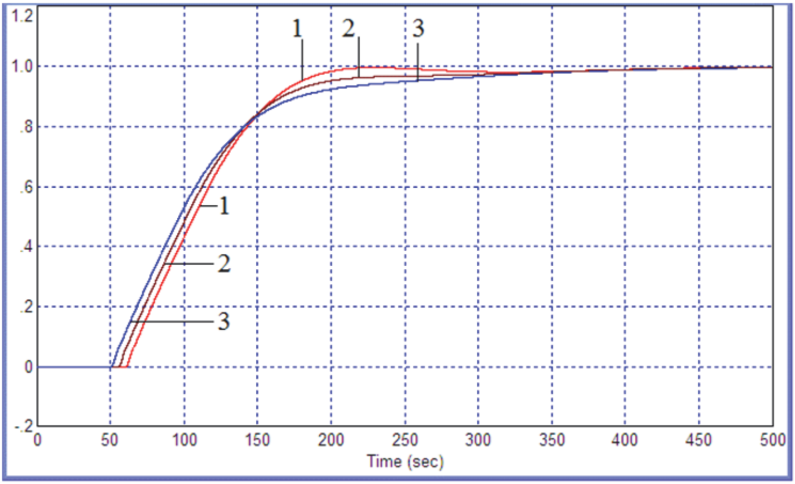


Fig. 2.7. The response of the system with the object (2.10) and PID regulator on a unit prescription step: line 3 – $\tau_1 = 50$ s; line 2 – $\tau_2 = 55$ s; line 1 – $\tau_3 = 60$ s

Object parameters may be changed in a wider range. Thus Fig. 2.8 shows process with the same regulator when time constant is changing from 20 s to 70 s. The system remains stable; the overshoot does not exceed 10% in the worst case. Such a system can be considered satisfactory for a number of technical problems.

Conclusion 5. At least in some cases it is possible to calculate regulator that the system remains operational with a significant change in any parameter of the object model within the limited range.

Example 8. To illustrate the statements that robust regulator is not optimal, let us calculate the optimal regulator for the object model (2.10) with the value of $\tau_0 = 20$ s. These coefficients are $K_P = 1.9$; $K_i = 0.0297$; $K_D = 12.86$. The resulting transient response is shown in Fig. 2.9 by line 2. If we give new meaning to the time constant $\tau_1 = 30$ s. The resulting graph is shown in Fig. 2.9 be line 1. In the first case, the transient process is of high quality: the process duration is small, not more than 100 *s*, the overshoot is negligible, less than 2%. In the second case, the overshoot increases to nearly 40%, the duration becomes 250 s. If further increasing the time constant, for example, to a value $\tau_2 = 40$ s, then the transient process in the system deteriorates even more. The corresponding graph is shown in Fig. 2.10; the overshoot is about 70%.

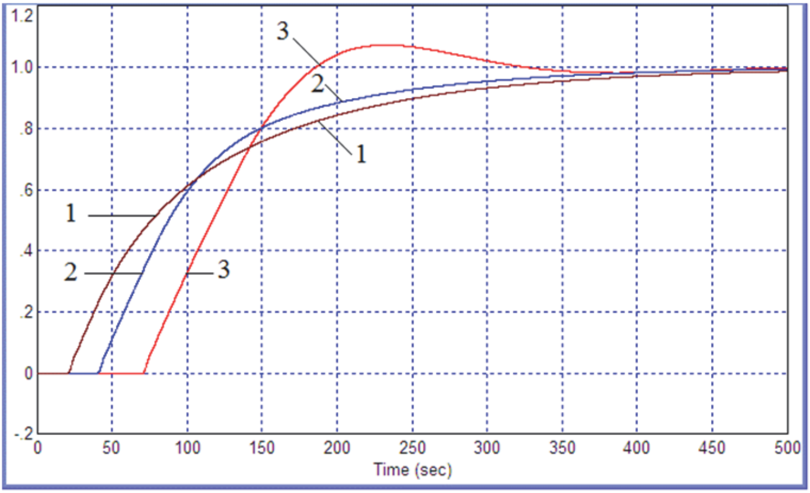


Fig. 2.8. The response of the system with the object (2.10) and PID regulator on unit step: curve 2 – $\tau_1 = 40$ *s*; curve 2 – $\tau_2 = 20$ *s*; curve 3 – $\tau_3 = 70$ *s*

Example 9. Next, let us calculate the optimal regulator for an object with a constant time $\tau_2 = 40$ s. The coefficients obtained are $K_P = 1.05$; $K_i = 0.0152$; $K_D = 12.45$. The transient process is shown in Fig. 2.11 with the line 2. It is also very attractive: overshoot of no more than 2%, duration about 150 s.

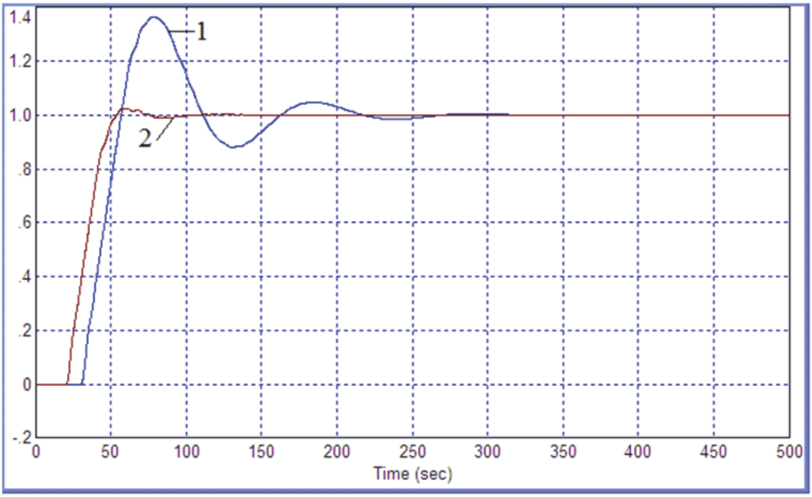


Fig. 2.9. The response of the system with the object (2.10) and the PID regulator of Example 8 at a unit step: line 2 – $\tau_0 = 20$ *s*; line 1 – $\tau_2 = 30$ *s*

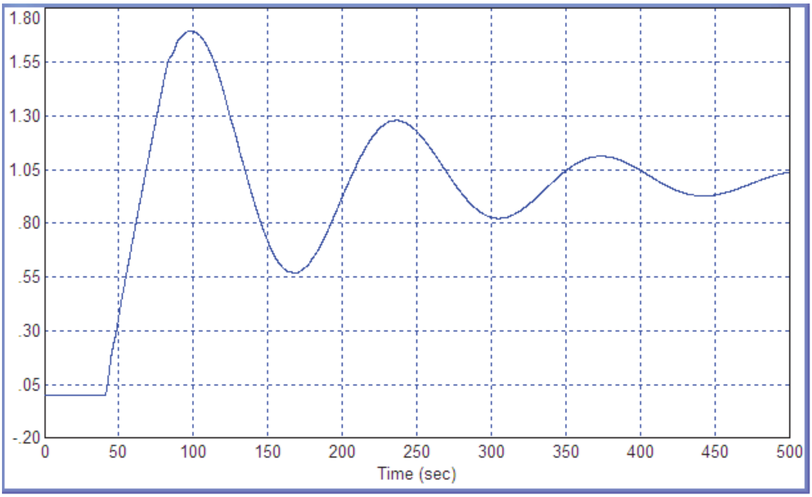


Fig. 2.10. The response of the system with the object (2.10) and PID regulator for Example 8 on the unit step when $\tau_3 = 40$ s

We return the value of the time constant $\tau_0 = 20$ s. The process is shown in Fig. 2.11 with line 1. It is not so attractive, since the duration is about 250 s. It is enough to compare this process with the process shown by the black line in Fig. 2.9 to see how much worse it is.

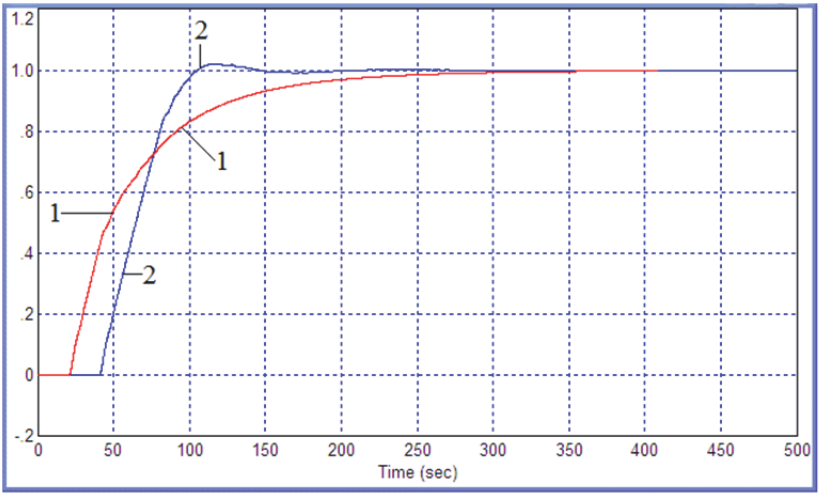


Fig. 2.11. The response of the system with the object (2.10) and PID regulator for Example 9 on the unit step: line 2 – $\tau_0 = 20$ s; line 2 – $\tau_2 = 30$ s

Conclusion 5. Robust controls are not optimal.
Conclusion 6. Optimal regulators are not robust.

The formulation of the problem of designing of adaptive system

That, what is claimed

1. The known model of the object or it can be determined as a result of the test cycle. It is also known that in this model, at least one of the coefficients vary in a certain limited range, the rate of change is limited.

2. There are requirements for a closed system.

That, what is required

To design the regulator, such that the system meets the requirements of claim 2 under conditions of claim 1.

Solving methods

The controller parameters cannot be changed during operation of the system. The ensemble of the system with the same regulators and with the different object model parameters is used for the optimization.

2.11. Advantages of digital regulators

Digital regulators for systems with negative feedback for the first time began to be widely used to control the slow objects, as well as to implement simple algorithms with low requirements on the quality and accuracy of the transient process (for example, relay control). Meanwhile, digital regulators can achieve the precision and broadband object control much better. Currently, cheaper and more compact regulators, characterized by a high stability, accuracy and reliability can be realized only by digital technology. Analogue technique is worse than digital one on all these parameters. Criteria of usability of digital technique is attainable speed of analog-to-digital converters (ADC) and their capacity, as the speed and capacity of digital-to-analog converters (DAC) are less critical and always higher than these parameters of ADC. Currently ADC with capacity of 14 bits or more and conversion frequency 200 MHz and above are available. It covers the needs of nearly all known real object control tasks. On the other hand, control of slow objects requires the implementation of large time constants that can be provided by digital technology without any restrictions, and analog technology - only through the use of large-capacity capacitor that causes the growth of the size and cost of the regulator. Thus, not only for controlling the slow objects and objects with a band of operating frequencies up to 100 MHz digital regulators have distinct advantages. For precision control, for control at low frequencies and (or) for control with interrupted feedback, only digital regulator can be used. The use of analog regulators in this case is impractical.

An advantage of digital regulators is also associated with the possibility of self-tuning and adaptation of the regulator parameters. Solving these problems with the only analog technology is unacceptable.

To simulate such systems *MATLAB* software, *Simulink*, *MathCAD*, *VisSim* and other automation software for design and analysis of locked complex dynamic systems can be used. We consider that the most adequate software for these purposes is *VisSim*. A careful analysis revealed possibilities of incorrect simulation results in *MATLAB*, *Simulink and MathCAD*. The reason for incorrect modeling is not so much in the lack of competence of the authors, as it is a lack of adequate software tools for this task, which consists of the excess accuracy of the implementation, which is not achieved in the real digital regulators. The signals that software *MATLAB* can calculate cannot be calculated with such accuracy by digital regulators. The digital regulator inevitably uses methods for numerical integration and derivation of the signals received in the form of a sequence of digital samples. According this algorithm *VisSim* software functions. Other software packages run by other algorithms, which cannot be used for digital regulators. Special options in *Simulink* are adequate too, but *VisSim* in principle cannot operate by another algorithm then algorithm of step-by-step calculating and this exclude mistake.

Comparative analysis revealed that simulation software *VisSim* 5.0/6.0 is the most adequate to the task of testing the developed algorithms, it is able to simulate the closed dynamical systems with linear and nonlinear elements, including objects with delay. Advantage of this software is performing of integration and derivation exactly in the same way as it can be performed on a microprocessor regulator, i.e. based on the available input signal samples, taken with a prescribed equal time interval. This software allows exploring of different methods of integration and derivation on the basis of which we can recommend the best method for implementation in a digital regulator. In contrast to *VisSim*, such software as *MATLAB*, *MathCAD* and others perform analytical calculation of derivatives and integrals that cannot be reproduced in practice by digital regulators, as analytical expressions of the input signals in reality are not available.

Note that in the simulation software *VisSim* it is necessary to choose one of several methods of integration. It is obvious that the implementation of digital control is also necessary to choose one of these methods. Studies have shown that in some cases the result varies with the method of integration. However, in realization of digital regulators, developers usually do not pay attention to this feature, considering that the integration and derivative, as such, give the desired result without depending on the method of implementation of these operations, while it is in not the case. Simulation program *VisSim* inevitably attracted developer attention to the problem of choosing of one of these methods. What is necessary is only to use the same method to regulator, which was used in the simulation.

Conclusion 7. Numerical optimization of digital regulators with the help of software *VisSim* provides the most adequate result, which can be repeated by regulators most accurately.

2.12. The two kind of tasks of regulator numerical optimization

We distinguish two different approaches to the problem of numerical optimization of a regulator. Regardless of whether an automatic, semi-automatic or manual optimization procedure is performed, the tasks can fundamentally differ depending on whether the research is conducted directly with the object, or with its model.

Accordingly, we divide two following classes of tasks.

Task 1. Numerical optimization of the regulator by experimenting directly with the object.

Task 2. Numerical optimization of the regulator by simulation using the object model.

When solving Task 1, the following difficulties can arise:

• **restriction on the mode**: inadmissibility of overshooting or exceeding of some pre-set level, inadmissibility of the oscillatory transient process, and so on;

• **restriction on the number of test cycles** or the high cost of each test;

• **low repeatability of experiments** due to noise, disturbances, non-stationary properties of certain parameters, and so on;

• **extremely low speed**, as a result of which many cycles of experiments are difficult or impossible.

A number of objects do not allow operation not only in unstable circuits, but even with insufficient quality of control, which leads to a damage of output products or to damage the object. The long duration of processes in real objects, taking into account the required number of simulation cycles (500–10,000), leads to an unacceptably long experiment time.

There are no said difficulties in solving of Task 2, but there are problems of adequacy of the used model to the real model of the object.

Specialists do not always adequately understand the importance of this problem. The fact is that in response to some specific effects, for example, on a stepwise action, it is possible to approximately build an object model and give an estimate of its parameters (the coefficients of this model). Modeling an object with the same effect demonstrates a fairly close coincidence of these experimental and model transient processes.

However, when implementing a system with a regulator, the type of transient processes in the system can be affected by those parameters of the real model that, during initial identification experiments, had an insignificant effect on their appearance in the open loop. The success of object identification should be evaluated not by coincidence of the object's response to a stepped or other test action in the absence of feedback, but by coincidence of the response of the system with the optimized regulator in its simulation and its real functioning in the presence of feedback. The results of this kind of coincidence are rarely given in the articles, which demonstrate the inadequate quality of controlled object identification in most practical cases.

This is the reason for the wide prevalence of the methods of practical solution of Task 1, the ineffectiveness of which is easily demonstrated with simulation. Namely: the methods of not damping oscillations, damping oscillations and others are specially designed to adjust the regulators based on actual tests of the system. Therefore, for example, the method of not damping oscillations suggests bringing the system to the state of not damping oscillations, to determine the coefficients of the regulator, leading to such oscillations and the frequency of the obtained oscillations. Based on these results, the regulator coefficients are calculated using the table, which is proposed to be used to control this object. Such methods are based on some simplified model of the object and on the relation between the above parameters. For example, it is most often assumed that an object has a model of the form of an aperiodic link and a element of pure delay consistently included with it. This hypothesis is often unfounded, therefore, the methods give far from optimal results, and often the result is simply not applicable. In addition, some objects simply cannot be brought to the state of not damping oscillations for technological reasons. Nevertheless, such methods are widely used in the practice of scientific research due to the lack of identification problems. In real practice, they are less common.

If the result is obtained on a real object, the problem of transferring it to practice does not arise. Any result, even far from optimal, in this case will be more reliable than the optimal result for not quite adequate model.

2.13. Objectives and tasks for the further research

The basis for further research is the thesis proved in [21–42]. It says that the toolkit for numerical optimization of regulators is most complete for solving the problems of design of regulators with a variety of features that complicate the solution of these problems. In the arsenal of this group of methods there are various cost functions, a number of effective regulator structures and the results of their comparative studies. Theoretical bases that confirm the correctness of simulation and the convergence of optimization algorithms are developed and recommendations for changing the cost function in the absence of convergence or in the case of unsatisfactory results of optimization are given.

Currently, research on improving the digital control of dynamic objects in a locked loop of automatic regulation is relevant. The flow of publications on this topic is not exhausted. Much attention is paid to these problems in the publications of journals “International Journal of Control”, “Technical Cybernetics”, “Automation and Telemechanics”, “Automation and Software Engineery”, and many others. Precision control is characterized by the requirements of reducing of the static error below 0.1%, and in some scientific problems (as, for example, in laser frequency standards), it must be below 0.001%, in which case rigid constraints are also imposed on the dynamic error tolerances. Satisfaction of these conditions can serve as a criterion for the effectiveness of the developed control algorithms.

Thus, to achieve this goal, it is necessary to solve the following tasks.

1. Development of methods and development of techniques for optimizing the control of technological objects with delay links, with nonlinear elements in their model, with output noise and output disturbance:

a) development of new effective regulator structures;

b) development of effective criteria for optimizing of locked systems;

c) development of effective optimization methods and investigation of their convergence.

2. Investigation of the possibilities of proposed methods and methods of object control under conditions of significant uncertainty in the coefficients of the object model and/or their random variation.

3. Development of methods for optimizing of the control of multichannel objects.

3. VISUAL SIMULATION OF UNLOCKED STUCTUTRES

3.1. The window of the software VisSim

As a tool for modeling, optimization and research of ACS, it is proposed to be the software *VisSim*. The window of this software (Fig. 3.1) contains the working field and panels of control commands and buttons (Fig. 3.2), whose purpose is intuitively clear.

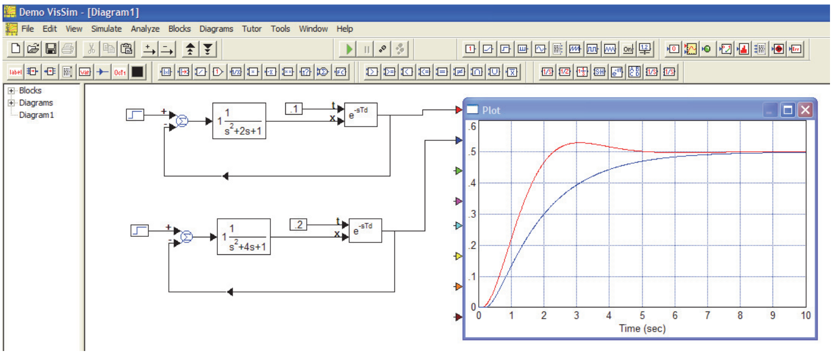


Fig. 3.1. A typical view of the VisSim software window

When mastering the software, it is recommended to investigate the purpose of each button independently. When a developer draws the marker "mouse" to any of the buttons there is an inscription-hint, explaining the purpose of the button. In addition to the buttons, commands from the command menu can be used to enter elements and set simulation parameters: *File, Edit, View, Simulation, Analysis, Blocks, Tools, Window,* and *Reference*.

The "File" menu allows you to create, open, insert or close a file, save, print, change properties, etc., and also open one of the last several editable files (see Fig. 3.3, *a*). Menu "Edit" allows you to cut, copy and paste fragments and make other changes (see Fig. 3.3, *b*). The "Simulation" menu allows you to configure simulation (optimization) or optimization parameters (see Fig. 3.4, *a*), and "Blocks" menu allows you to select the standard blocks from the proposed list (see Fig. 3.4, *b*).

The first 7 menu buttons (see Fig. 3.2) are understandable to users of personal computers without additional explanations.

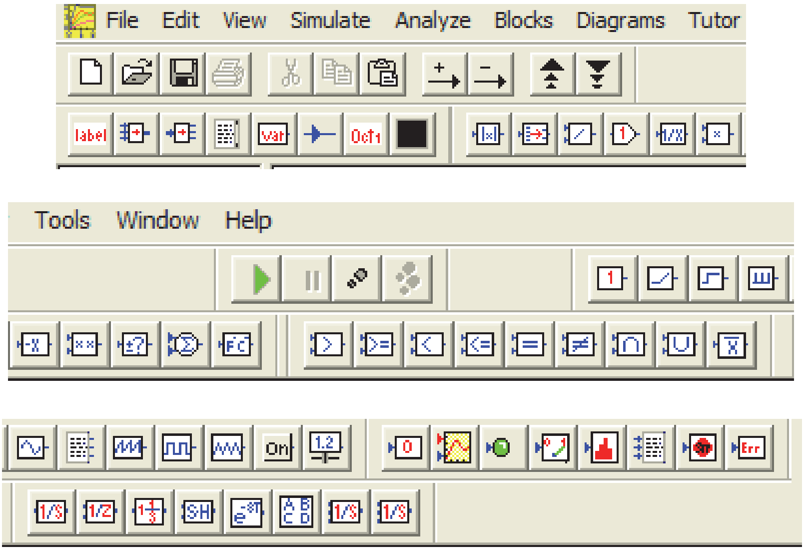
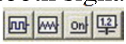


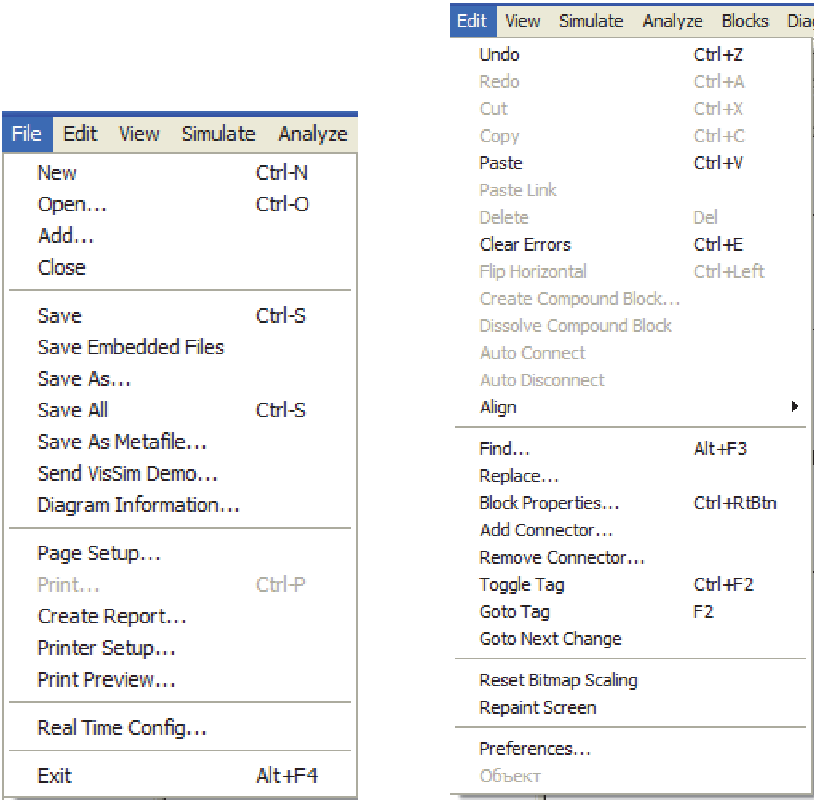
Fig. 3.2. VisSim control buttons

Buttons  and  allows you to add or remove inputs from other units. Buttons  and  also allow you to increase or decrease the zoom level. The button  starts the simulation or optimization procedure, depending on the selected mode.

Buttons for inserting blocks allow you to insert the corresponding frequently occurring blocks. So, with the help of buttons , you can insert respectively a constant-forming unit, a linear-rising signal generation unit, a step-jump forming unit (unlike a constant block in this block, you can set the delay amount), a comb-forming unit, a harmonic signal generating unit, block for data import from the file and the sawtooth signal generation unit.

Buttons  allow you to insert blocks of formation of rectangular signals and triangular signals, a block for switching on and off the simulation, and a "potentiometer" block (that is, a constant driver whose value is controlled by the position of the slide rheostat).

The buttons  allow you to insert display units: a terminal (to indicate the value of the signal at the end of the simulation), an oscilloscope (for displaying 1 to 8 graphs as a time function or X-Y axes), a probe indicator, an indicator in the form of a pointer, indicator in the form of a histogram, a block of export to a file, a block "stop" and an "error" block.



a *b*

Fig. 3.3. The menu commands "File" (*a*) and "Edit" (*b*)

The button  allows inserting a fragment of the bus with a fixed position on the graph, which can be arbitrarily changed by dragging this fragment along the field. This makes it possible to fix the location of the tire, which provides a more graphic representation of the project.

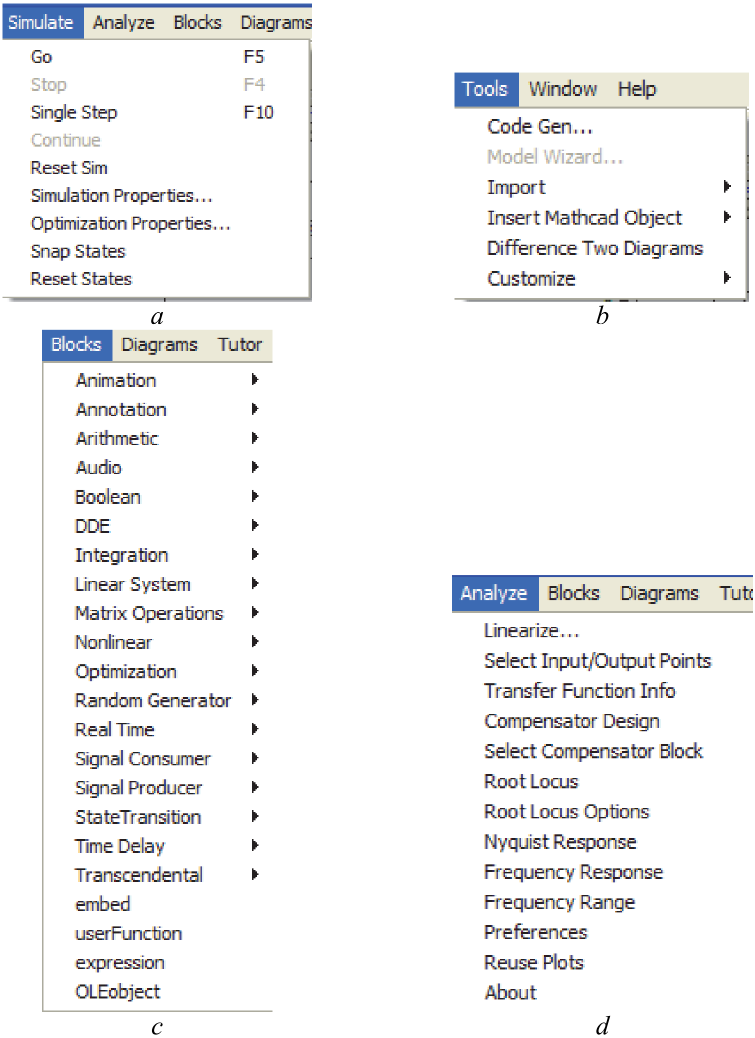


Fig. 3.4. The menu commands "Simulate" (*a*), "Tools" (*b*) and "Blocks" (*c*) and "Analyse" (*d*)

The resulting "line fragment" block can be turned inversely, like most other blocks. To rotate it, just select the block and click the right mouse button to bring up the menu, which includes the "reverse" function. In the images of the block, by default, the inputs are located on the left, and the outputs are on the right. This also applies to units that have only inputs or only outputs.

The button  allows you to insert the rectifier block (forming absolute value of the input signal), and the button  allows inserting of the amplifier (multiplier by the coefficient). By default, this factor is equal to one, to change it, simply call the editing menu with the right “mouse” button. It is permissible to use the letter designation of the coefficient if the value of this letter variable is numerically determined throughout the simulation.

The button  allows you to insert a name for the line. After that, you can assign any alphabetic or alphanumeric name in editing mode, for example, X, Y, e1, V2, etc. It should be remembered that small and capital letters are treated as different letters, for example, the names A2 and a2 will be perceived as different names of lines. Equally named lines are considered to be the same line. A block  with a fixed line name can be found in the project as many times as desired, all its outputs will mean the same point. In this case, only one time its input can be use to connect some signal which may be sent from the output of another block. An attempt to send any signal to the input of another block with the same name will be rejected by the program as an error, which is natural, since the outputs of different blocks can not be connected with each other. The inputs of any number of blocks can be connected to each other without restrictions, but all together they must be connected to only one output of the only block. In this sense, there is a complete analogy to the electrical circuits, with voltage generators (unlike them, current sources can be combined in outputs, but the *VisSim* program does not provide this). Buttons  allow you to enter an integrator (of two kinds), a linear filter described by the transfer function, a sample-and-hold device and a delay link, respectively.

Acquaintance with other buttons occurs as they are used. For a more detailed reading, it is recommended to study information from the Klinachev website [19–20].

3.2. Start of the working in the software

When you open the program, the working window occurs, as Fig. 3.5 shows. The top line contains commands, below line contains the buttons for working with files and for adding, removing or fixing elements in the project.

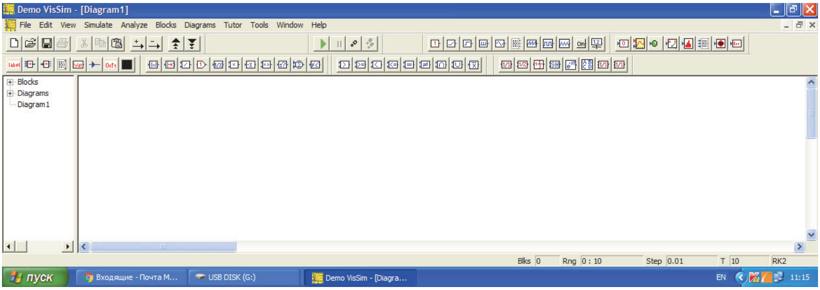


Fig. 3.5. The *VisSim* program window

To start modeling, it is necessary to create a model of the system with the display device. Thus, we need a signal source, a device that converts these signals and an indicator of the received signals. In the simplest form, for example, it can be a harmonic signal generator, a nonlinear element, and an oscilloscope.

An example of such a structure for modeling and the result is shown in Fig. 3.6. It uses two consecutive nonlinear devices, namely the dead band element and the limiter, included in series. Note that if the nonlinear elements are interchanged, the result (output signal) will change. Fig. 3.7 shows the graph that is obtained with the indicated change in the order of the nonlinear elements. The amplitude of the received signal has noticeably decreased.

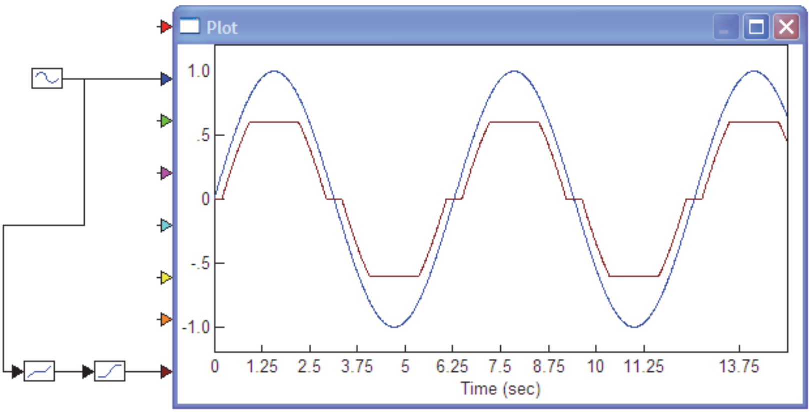


Fig. 3.6. Modeling of the sine signal generator and non-linear transforming elements

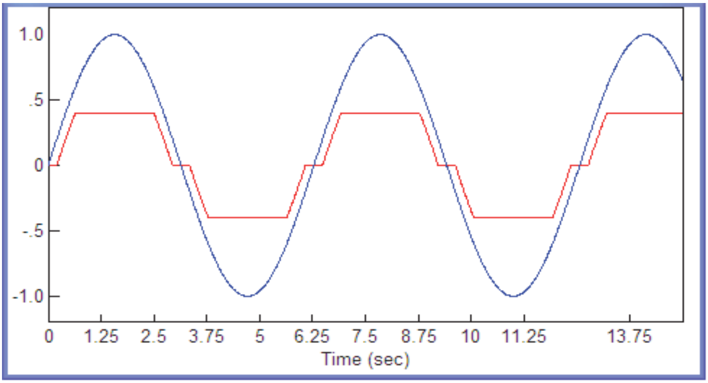


Fig. 3.7. Modeling of the sine signal and non-linear transforming elements after changing of the order of these elements

3.3. Tuning of the parameters of the simulation and optimization

For tuning of the parameters of simulation (mathematical modeling), it is necessary to call up the settings window (Fig. 3.8).

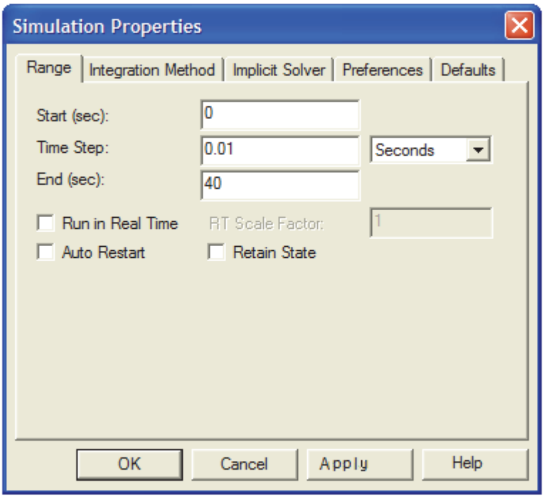


Fig. 3.8. The simulation settings menu

By default, the simulation starts at the zero point in time. At this point, all the constants jump to their prescribed value. The values of step functions jump at the same moment to their prescribed values, if they did not have nonzero delay value. The time step (or step of integration) determines the magnitude of the quantum in time, i. e. the period of time sampling in the simulation. The end of simulation time specifies when the simulation is complete. Used can also select the units of time (seconds, minutes, hours). The software allows realization of real-time mode by ticking the corresponding menu window. In this case, the simulation will occur at the exact tempo that corresponds to the selected time units and with the selected integration step. This option is only useful if the software is applied to generate real signals that interact with real external devices, for which input and output port is available. In the case of modeling with the goal to determine the stability or to optimize the regulator, there is no need for a complete correspondence of the computer time to the real one. On the contrary, such a correspondence is undesirable, otherwise the optimization of the regulator for an object described by time constants of the order of several minutes would take (for example, 500 iteration steps) about several thousand minutes, i. e., tens of hours. For the same reason, one should not attach special importance to units of time. Any unit of time can be treated as "second", depending on the speed of the objects it can be microsecond, millisecond, minute, hour, day, month, year, etc.

The operator always has the possibility to apply the time scaling method, considering the conditional "second" is such time unit at which most model and controller coefficients related to time take the most middle values, i. e. their logarithm is closest to unity. Indeed, it is inconvenient to deal with too small or too large values, especially for a large order of an object, taking into account that each polynomial coefficient (in the numerator or denominator of the transfer function) contains the corresponding scale (in time) coefficient to the same power as the argument s.

Example 10. Consider the transfer function of an object:

$$W_O(s) = \frac{500 + 750s}{20 + 50s + 1000s^2}. \tag{3.1}$$

It can be reduced it to the canonical form, ensuring the unit values of the free term in the numerator and denominator polynomials:

$$W_O(s) = \frac{25(1+1,5s)}{(1+2,5s+50s^2)}. \qquad (3.2)$$

In this form, this transfer function can be used for simulation more easily, since the available coefficients not differs too much (in order of magnitude) from unity.

To place the model of the transfer function in the form of a rational fraction, it is necessary to choose “blocks - linear systems - transfer function” or press the button with the symbol $\boxed{1\frac{1}{1}}$.

A window for setting up a simulation of the transfer function in the form of a rational fraction (3.2) is shown in Fig. 3.9. Fig. 3.10 shows the result of modeling the response of an object with this transfer function to a step unit input action.

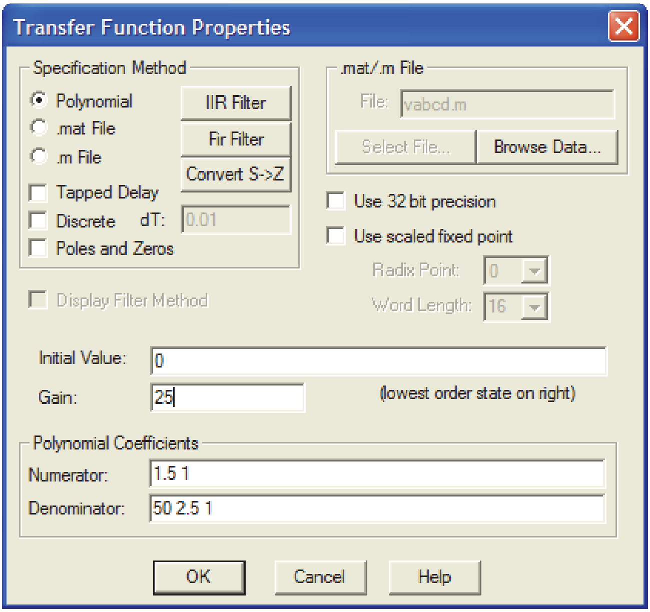


Fig. 3.9. The menu for setting the parameters of the transfer function (3.2) in the form of a rational fraction

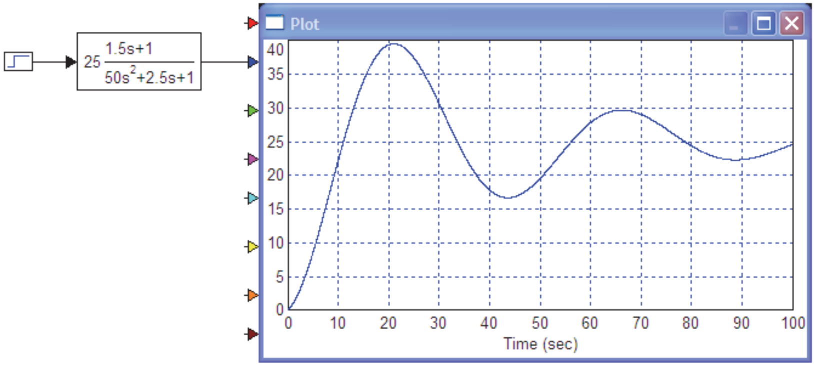


Fig. 3.10. The result of modeling of the object response with transfer function (3.2) to a step unit input

Example 11. Consider the transfer function of an object:

$$W_O(s) = \frac{50}{1 + 0{,}001s + 0{,}000001s^2}\,. \tag{3.3}$$

Replacing the seconds by milliseconds, we get the transfer function of a simpler form:

$$W_O(s) = \frac{50}{1 + s + s^2}\,. \tag{3.4}$$

This transfer function is more convenient for modeling, but now it should be taken into account that the simulation is performed not in seconds, but in milliseconds, that is, a factor of 1000 in relation to frequencies and factor 0.001 in relation to time are introduced. This factor should be taken into account when implementing the regulator.

3.4. Choice of the time sampling step

To ensure the best simulation accuracy, the integration step should be chosen as small as possible. Fig. 3.9 shows an example of the result of simulating of the harmonic transient process for different values of the time step. It is seen that the form of the graph is distorted for too large step is, and the distortions are negligible for sufficiently small time step.

From the standpoint of saving the computational resource and the calculation time, one should strive to choose the reasonable value of this step, not too small. Especially acute is the problem of its choice in solving the optimization problem, since in this case of too small time step it can lead to too much computation time.

The first rule. The step must be small enough that the increment of all signals in the final version of the simulated system is negligible during the transient process (in comparison with the magnitude of the transient process, provided it is stable). The magnitude of the transient process of an unstable system can be not taken into account. Therefore, it is permissible if in intermediate stages this step is not sufficiently small, but in the final stages of modeling it should be chosen correctly. Worse, if the situation is reversed: at first the step it will be correct, but at the final stage due to the increase in the gain coefficient the system's speed will increase, and the previously chosen step will become incorrect. Therefore, the verification of the correctness of the selection of the time step must be repeated at the final stage.

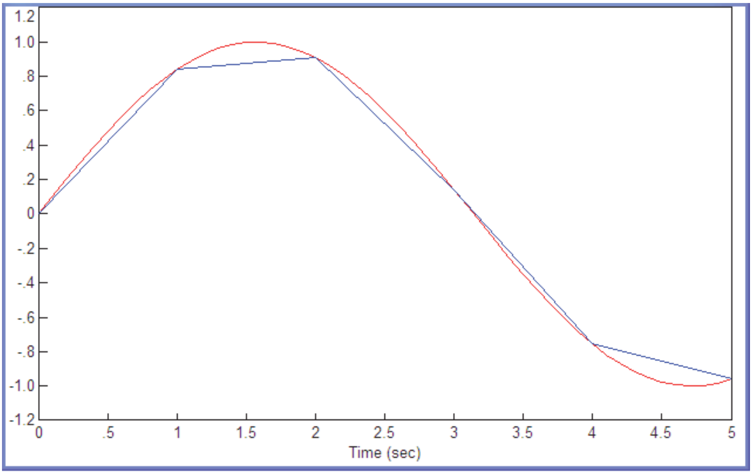


Fig. 3.9. Graphs of transient processes for various settings of the time step (0.1 *s* and 1 *s*)

The second rule for choosing a step can be formulated in this way: the step should be 20–100 times smaller than the smallest time constant in the object and regulator divided by the gain in this loop.

Since the gain in the loop during the regulator optimization process can vary, as well as the regulator parameters, it can be assumed that even if in the initial stage of the simulation the step was chosen quite justifiably and correctly, after completing the calculations with the new values of the regulator parameters, this step may turn out to be incorrect, i. e. too large.

The third rule is necessary and sufficient. It states: if after the simulation the step is reduced by 2–4 times, after which the repeated simulation gives practically the same results, the previously chosen step was sufficiently correct.

3.5. Choice of the integrating method

When choosing the method of integration, a series of different methods are suggested, including *Euler* with lag, trapezoidal, *Runge-Kutt*, different orders, etc. (see Fig. 3.10).

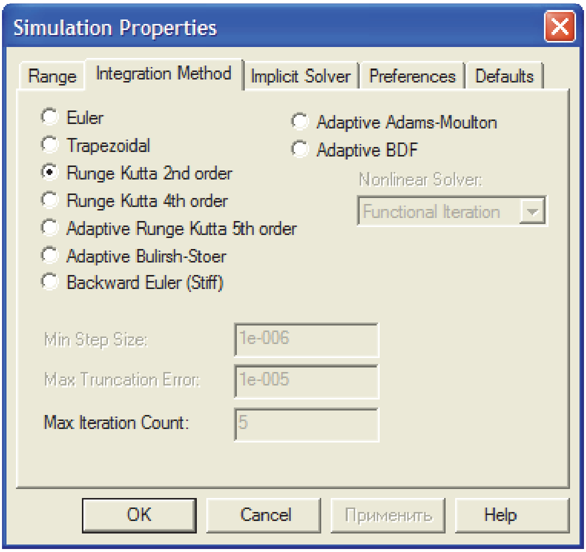


Fig. 3.10. Integration method setup menu

We recommend choosing the *Euler* method with delay (Euler's simple method) or the *Adaptive Bulirsh-Stoyer* method (ABS), depending on the characteristics of the problem.

The rationale for choosing the *Euler* method is given in. To substantiate the influence of the choice of the integration method on the correspondence of the results of the sequential application of mutually inverse operations to the characteristic signals is investigated. If the test signal is sequentially transmitted through an equal number of derivative and integrating links (in any order), then the output signal should be the original signal. This is not always achieved, and not always with every choice of the method of integration. When choosing a simple *Euler* method and providing a sufficiently small integration step, the best match is achieved. Therefore, we recommend a simple Euler method. In this method, the integral of a function is calculated as the sum of rectangles whose base is the intervals equal to the step of integration, and the height is the values of the function at the beginning of this step. Since the integration step is sufficiently small, the computation of the integral is sufficiently accurate.

The rationale for choosing the ABS method is justified by the fact that only with this choice the open model of the object in the form of many loops with feedbacks is adequately modeled.

Let consider the transfer function of the form (3.4). The relation between the input $U(t)$ and the output signal $X(t)$ is described in this case by an operator-form equation (in the Laplace transform domain):

$$X(s) = \frac{50}{1 + s + s^2} U(s) . \tag{3.5}$$

This relationship can be rewritten in the following form:

$$X(s)(1 + s + s^2) = 50U(s) . \tag{3.6}$$

Also it can be rewritten in the following form:

$$s^2 X(s) = -sX(s) - X(s) + 50U(s) . \tag{3.7}$$

Hence, it is easy to obtain the following relation:

$$X(s) = \frac{1}{s}\left\{-sX(s) + \frac{1}{s}[-X(s) + 50U(s)]\right\} . \tag{3.8}$$

Fig. 3.11 shows the structure for modeling the system by the relation (3.5) and by the relation (3.8). Transient processes approximately coincide. However, if we consider the difference of these processes, as shown in Fig. 3.12, the difference of these graphs reaches a value of 0.2 units, that is, slightly less than half a percent. When modeling by different methods, this difference assumes different values, in particular, for this example in the

Bulirsh-Stoer simulation, this peak error value reaches $1 \cdot 10^{-6}$, and the minimum value in the peak reaches $-3.5 \cdot 10^{-6}$.

When modeling the fourth-order *Runge-Kutta* method and the adaptive *Runge-Kutta* method of the 5th order, this error value in the peak reaches a value of $3 \cdot 10^{-9}$, and the minimum value at the peak is -10^{-9}.

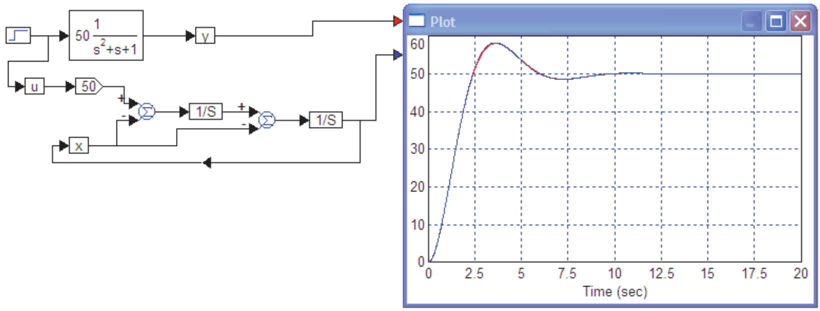


Fig. 3.11. The structure for modeling objects by models (3.5) and (3.8)

Thus, in this case the best coincidence is achieved by *Runge-Kutta* methods of the 4th or 5th order. An error in modeling by the second-order *Runge-Kutta* method is incommensurably greater. In more complex structures, in some cases, the smallest error was obtained by modeling with the *Bulirsh-Stoer* method.

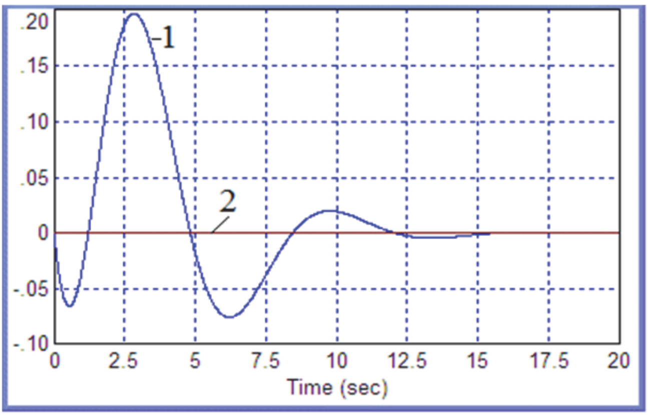


Fig. 3.12. The difference in the graphs in Fig. 3.11 for Euler simulation (line 1) and ABS (line 2)

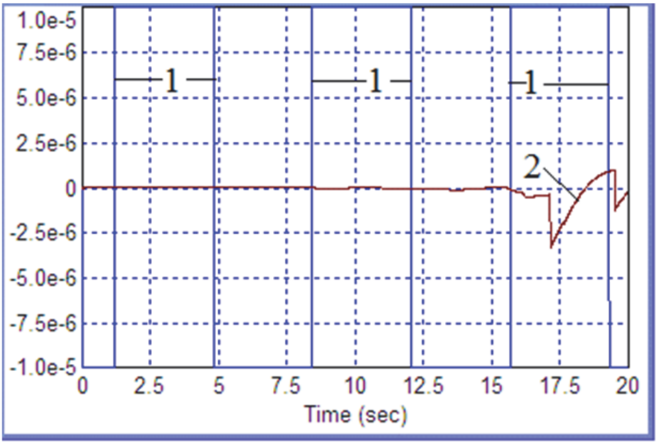


Fig. 3.13. The same as in Fig. 3.12 with an enlarged scale along the ordinate axis

Conclusion 7. The choice of the method of integration depends on the problem being solved. Among the most appropriate methods, one can recommend the simple *Euler* method, the *Runge-Kutta* method of the 4th and 5th order, and the *Bulirsch-Stoer* method.

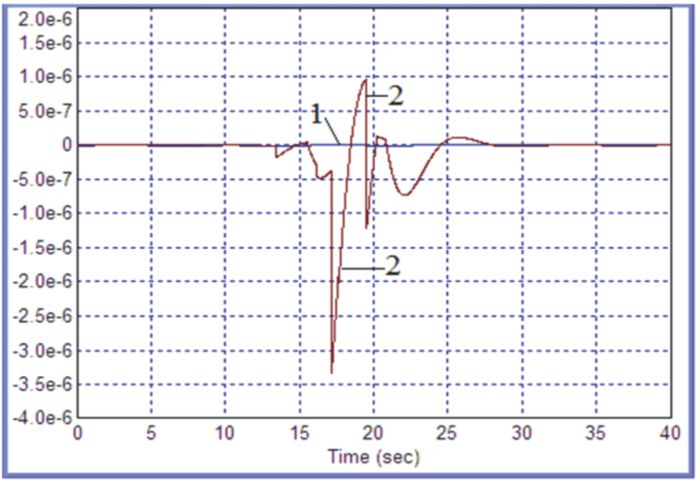


Fig. 3.14. The difference in the graphs in Fig. 3.11 for 4th-order *Runge-Kutta* simulation (line 1) and ABS (line 2)

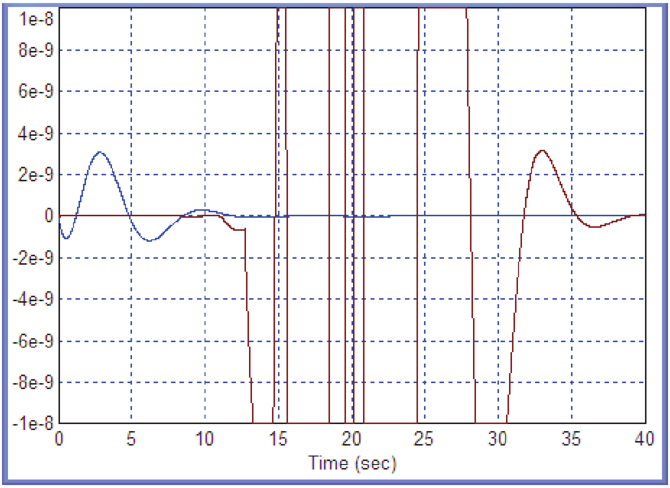


Fig. 3.15. The same as in *Fig*. 3.14 at an enlarged scale along the ordinate axis

In a number of problems, only one of these methods gives adequate result.

3.6. Tuning of the parameters of simulation and optimization

To configure the parameters, you must choose one of three optimization methods: *Powell, Polak-Ribiere or Fletcher-Reeves*. *Fig*. 3.16 shows the corresponding configuration window. It is also possible to specify user's own method (custom). The difference in the action of the optimization methods is investigated by our own experience. From the point of view of the problem of optimization of regulators, none of the three built-in methods have clear advantages in relation to others. With the correct choice of the objective function, any of them is effective enough to achieve its minimum, which is the goal of optimization. The number of optimization steps, as a rule, is excessively high. When modeling with a virtual object (but not with real and not in real time) this is not such a significant drawback.

The software implements two types of optimization: reduction of the objective function "Constraint" to zero and achievement of the minimum of the "Cost function". The last of the listed types is the most relevant for optimizing the regulator parameters. To run this mode, you must tick the “Perform optimization” window. The “Max Iterations” window sets the maximum number of iteration steps. Theoretically, after the set number of steps is completed, the program must stop, indicating the best result achieved for the specified number of steps. In practice, all the versions examined ignore this limitation, in which one can easily verify by putting a small number of iterations and observing the program's actions. The “Error Tolerance” window allows you to set the tolerance for the error in finding the minimum of the extremum of the objective function. This value can be set much less than suggested by default, since the proposed value of “1” may be too rough, if the objective function takes not too much significance.

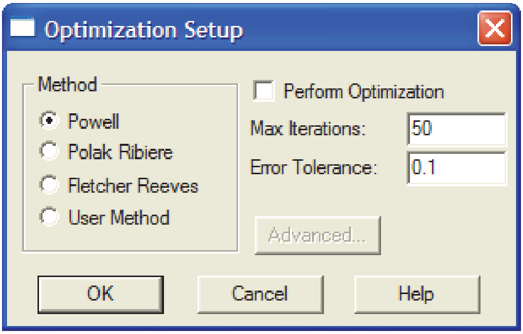


Fig. 3.16. Optimization settings menu

3.7. Simulation of the response of the linear link

To simulate the response of the link described by the linear transfer function, it is enough to enter such a link, the generator of the corresponding test signal and the oscilloscope, and then launch the button “Play” (Start).

Example 12. We define a transfer function of the form

$$W(s) = \frac{1}{1 + 2s + s^2}.$$

For this purpose, when editing the created block, we set its parameters, as shown in Fig. 3.17.

After starting the simulation, a transient process will be obtained, which is shown in Fig. 3.18. The shape of the oscilloscope window can be changed by editing. Pressing the right button on the oscilloscope block calls the window shown in Fig. 3.19.

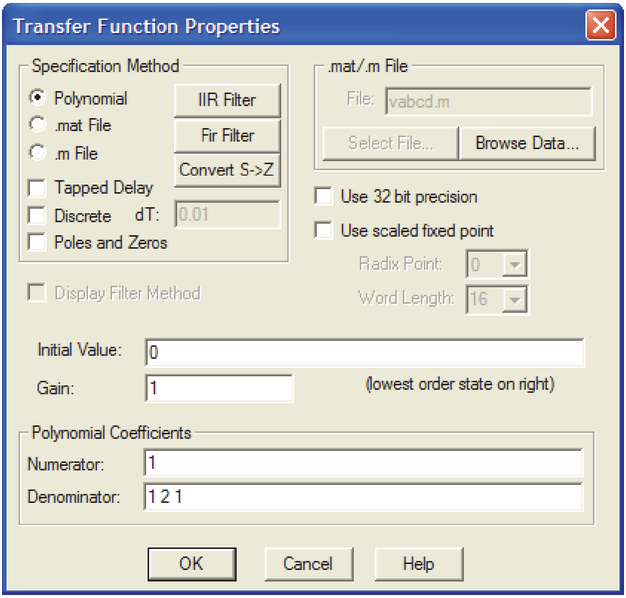


Fig. 3.17. Settings menu of the transfer function block

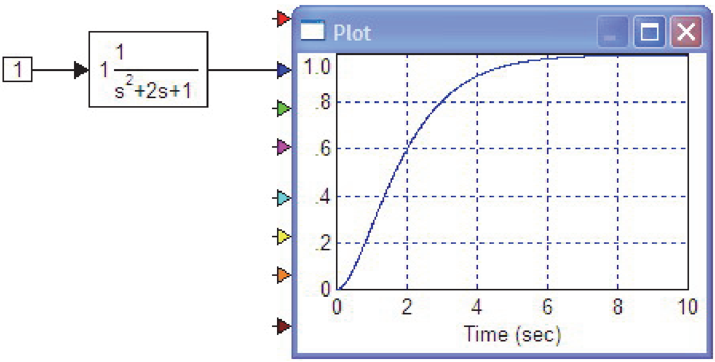


Fig. 3.19. Modeling of the response of the transfer function block

Having selected the option “Over Plot” and the number “4” (the maximum allowed number is 8), we obtain simultaneous display of four graphs. You can change, for example, the coefficient of the first power of *s* (see Fig. 3.20).

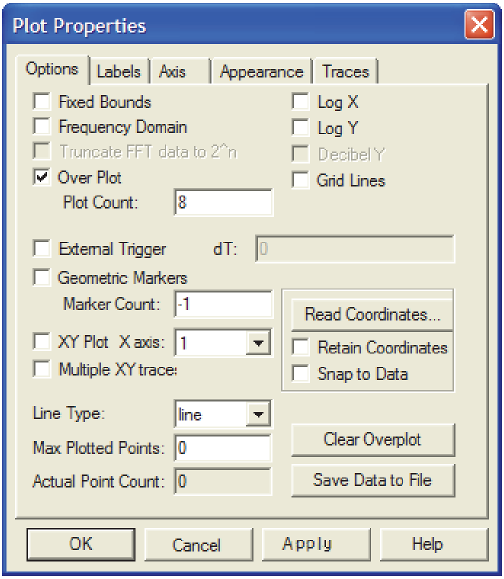
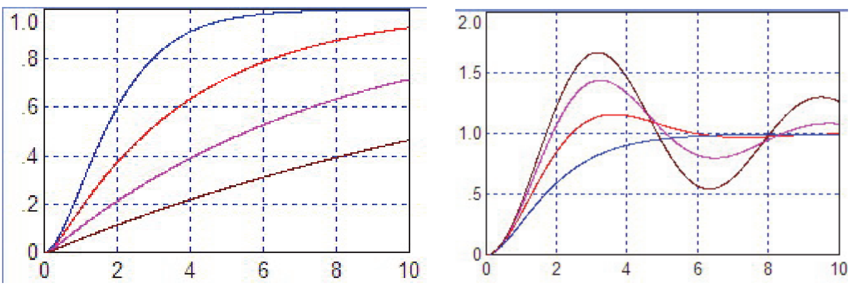


Fig. 3.19. Oscilloscope parameters editing menu



a *b*

Fig. 3.20. Dependence of the transient process on the coefficient at *s*: *a* – an increase of it in 2, 4, 8 and 16 times; *b* – decrease in 2, 4, 8 and 16 times

3.8. Simulation of the response of non-linear links

When modeling the response of nonlinear structures, it is important to follow the order of inclusion of nonlinear and linear links. Changing the order of the sequence leads to changes in the result.

Example 13. Transferring the link of the restriction from the first place to the second and third in the example of *Fig*. 3.21 causes a corresponding change in the response of the entire chain of these elements. It is easy to see that a change in the order of succession of linear links within a chain containing only linear links does not cause any changes in the result beyond the limits of this chain.

It should be remembered that in practice, there are no ideal linear links; each element of the system has some form of nonlinearity (and sometimes several). At a minimum, each element is characterized by a limitation of the input effect. Therefore, in practice, most often the order of the elements is very significant. For the same reason, in modeling, one should try to take into account all the nonlinear elements entering the real system, and strictly follow the order of their following.

In an analytic study of the system, taking into account all the nonlinearities unnecessarily complicates the system of equations, which, as a rule, is extremely complicated for theoretical analysis. In modeling, the account of nonlinearities is not particularly difficult in connection with the presence of a sufficiently wide set of characteristic nonlinearities and with the possibility of specifying their parameters.

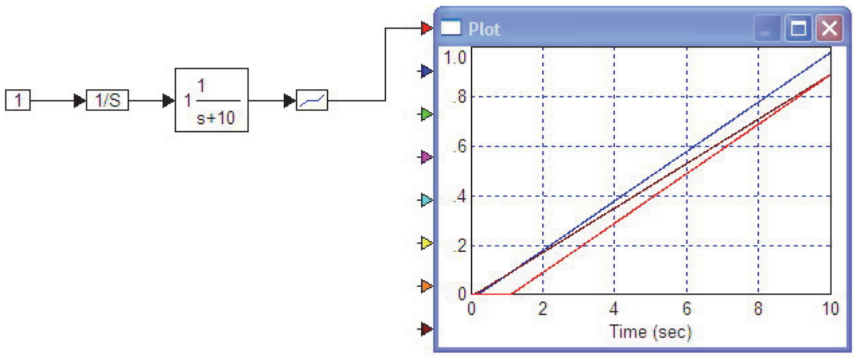


Fig. 3.21. Dependence of the transient process on the order of the elements in the case of the presence of a nonlinear element

Thus, in the arsenal of nonlinear blocks there are blocks: a multiplexer, a coincidence detector (intersection, cross detection), a dead band, an all-part selection, a limiter, a tabular nonlinearity, a high-end block and a low-level pick-up unit, a switch, a quantizer, storage. In addition, there are blocks that perform arithmetic operations (mark extraction, rectification, signal multiplication, signal raising, etc.).

4. VISUAL SIMULATION OF THE LOCKED STRUCTURES

4.1. Simulation of the locked linear system

Visual simulation of locked ACS consists of obtaining graphs of output signals of locked structures under the influence of given signals on the inputs of these structures. In addition to obtaining graphs, it is also possible to obtain the corresponding frequency characteristics of locked structures. The resulting graphs can be saved to a file for later processing. In addition, the simulation makes it possible to clarify the dependence of the output signals on the change of certain parameters of individual elements in the form of a family of graphs.

Fig. 4.1 shows an example of the result of modeling of a locked system. Since in traditional regulator structures feedback is generally negative and single, it is these structures that are further demonstrated without additional justification.

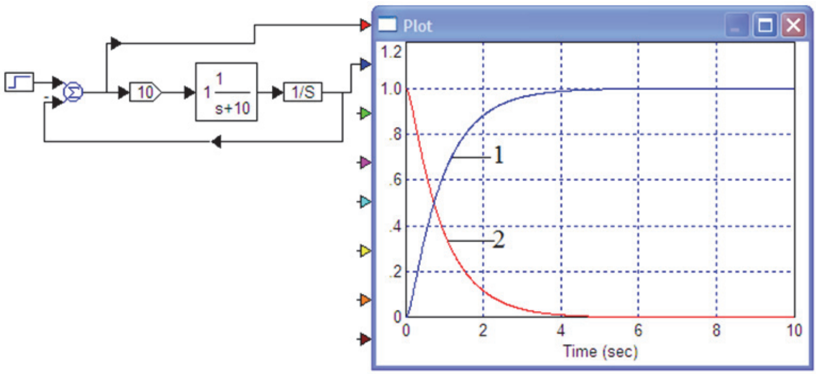


Fig. 4.1. Simulation of the response of the system to a jump in the control action (line 1 is output signal, line 2 is error in the system)

As a rule, any real system is affected by disturbance, which cannot be measured, but influences to the output signal. In the simulation, such an effect is formed as an additive signal on the output of the object, as shown in Fig. 4.2.

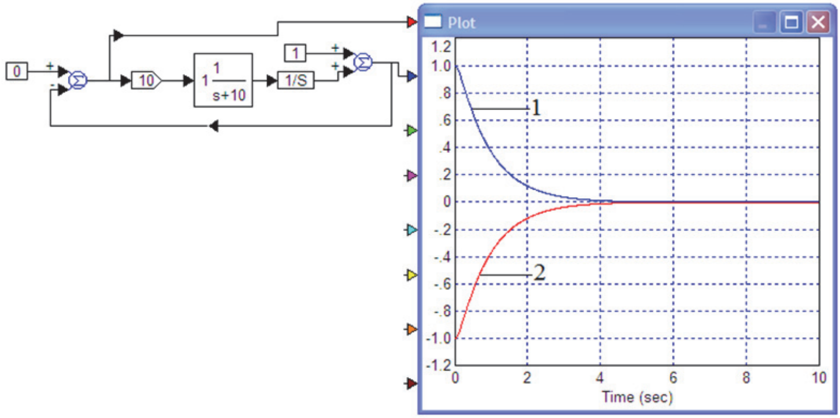


Fig. 4.2. Modeling the response of the system to the jump of the disturbance (line 1 is output signal, line 2 is error in the system)

The task of designing a regulator is to find a mathematical model of it that provides the most accurate repetition by the object output signal $Y(t)$ of the reference signal $V(t)$ with the least possible dependence of this signal on the disturbance effect $H(t)$.

It is evident by comparing the graphs of Fig. 4.1 and 4.2, that the response of the system to a step unit jump is interrelated.

The image of the output signal $Y(\omega)$ is related to the images of the target $V(\omega)$ and the interference $H(\omega)$ by the following relation:

$$Y(\omega) = \frac{W(\omega)}{1+W(\omega)} V(\omega) + \frac{1}{1+W(\omega)} H(\omega) .$$

Here $W(\omega)$ is the transfer function of the open loop, equal to the product of all transfer functions of the loop. Based on this, we can write the relations for the output signals of the structures in Fig. 4.1 and 4.2:

$$Y_1(\omega) = \frac{W(\omega)}{1+W(\omega)} V_1(\omega) , \quad Y_2(\omega) = \frac{1}{1+W(\omega)} H_2(\omega) . \qquad (4.1)$$

Here, $Y_1(\omega)$ is the transform of the output response signal $y_1(t)$ to the step unit jump of the prescription $v_1(t)$, $Y_2(\omega)$ is the transform of the response output signal $y_2(t)$ to the same jump of noise $h_2(t)$ (the index refers to the number of the simulation scheme). For $V_1(\omega) = H_2(\omega) = \sigma(t)$ (the symbol $\sigma(t)$ denotes the unit step jump at the instant $t = 0$), we obtain $Y_1(\omega) + Y_2(\omega) = \Sigma(\omega)$, that is, $y_1(t) + y_2(t) = \sigma(t)$, where $\Sigma(\omega)$) is the transform of the step function $\sigma(t)$. Thus, having the result shown in Fig. 4.1, we can easily calculate the result shown in Fig. 4.2, and vice versa. That is, in the case of a linear single-loop system with a single feedback, for simplicity of analysis, the system response to the perturbation $h_2(t)$ can be not considered separately, since the response to the control action $v_1(t)$ is exhaustively informative. Any of these responces is sufficient. This situation is not similarin the case of nonlinear systems, as well as for systems with many circuits and to systems that use non-unit feedback or (and) in the direct path between the reference and the subtractor there is an element with a non-unit transfer function.

Example 14. Consider an object of the first order. Control of such an object can be provided by a proportional regulator (Fig. 4.3). An increase in the proportional regulator coefficient simultaneously ensures an increase in speed and a reduction in the static error. Fig. 4.4 demonstrates that further increase in the gain factor gives the same effect without causing a stability violation (if the integration step with each simulation is correlated with the process rate). If the integration step is not changed, then the stability of the system will be violated, but such modeling would be not correct.

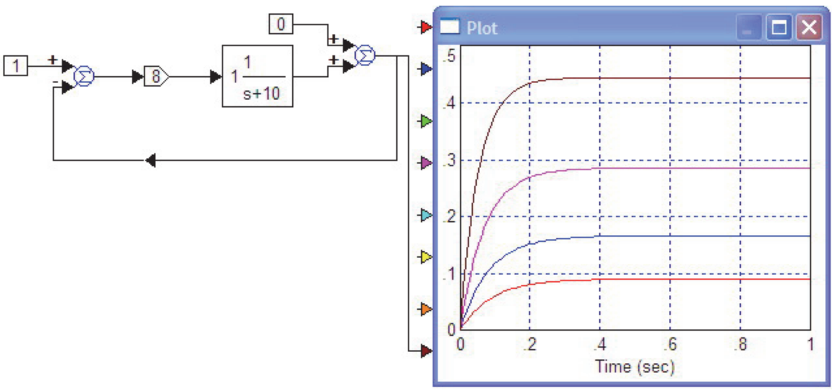


Fig. 4.3. Transient processes when the coefficient of proportional regulator varies from 1 to 8

Conclusion 8. A first-order system with negative feedback is stable with any gain.

This result follows from the theory of automatic control, Example 14 only illustrates it.

If the stability is still violated when modeling a system with a first-order object and a proportional regulator, this will indicate a violation of the conditions for the correctness of the simulation, namely: the speed of the system will become so large that the previously chosen integration step is already insufficiently small.

Conclusion 9. A violation of the stability of a first-order system with an increase in the coefficient indicates an incorrect simulation.

This statement follows from theory and it does not require experimental or model evidence, but those who wish can be convinced of its validity by modeling.

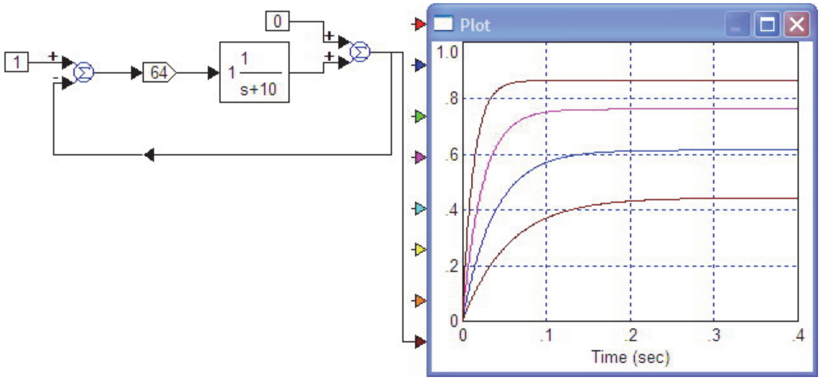


Fig. 4.4. Transient processes when the coefficient of the proportional regulator varies from 16 to 64

Example 15. Consider, under the same conditions, a system with a second-order object (see Fig. 4.5-4.7). As the proportional regulator K coefficient increases, the response time also increases, and the static error decreases. However, after reaching a certain value of this coefficient, an overshoot occurs, which then increases with increasing of K. Over 100% overshoot almost always means a loss of stability. Very rarely such overshoot occurs in a stable system. An ideal second-order system can be on the stability boundary, but formally not beyond this boundary. In such a system, there are no steadily increasing oscillations, although not damped oscillations may occur. In this case, it is sufficient to have one more element in the system with the slightest additional inertia, so that the system becomes unstable.

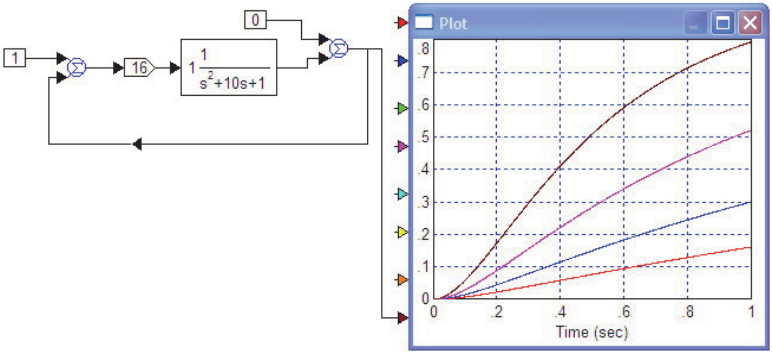


Fig. 4.5. Transient processes with a second-order object when the proportional regulator coefficient varies from 1 to 16

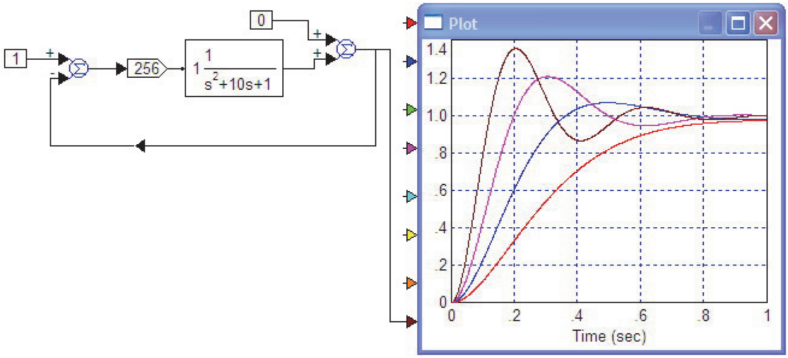


Fig. 4.6. Transient processes with a second-order object when the proportional regulator coefficient varies from 32 to 256

In practice, there are no ideal objects of the first or second order. Really its order is much higher. Therefore, any practical example of a system originally described by a first- or second-order model has not full and correct mathematical model. Hence in practice even such object can be brought to the loss of stability by increasing the gain. In software *VisSim*, additional inertia gives the difference of numerical integration from the ideal mathematical performance of this operation. Therefore, when modeling a second-order system in this program, an unstable process can be obtained. Thus, modeling in this program is more similar to the practical situation than theoretical analysis or modeling, for example, in software MATLAB.

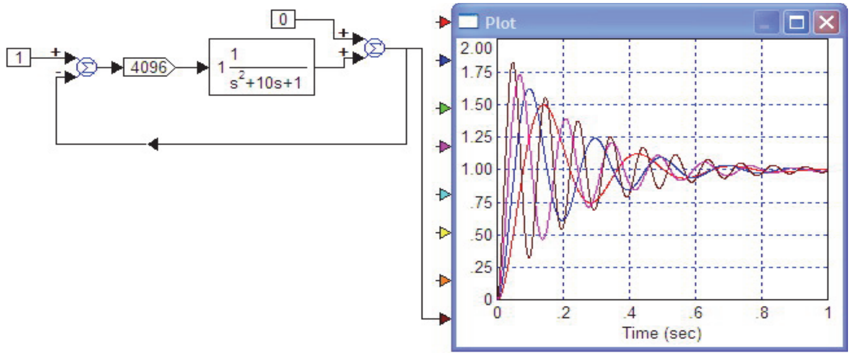


Fig. 4.7. Transient processes with a second-order object when the proportional regulator coefficient varies from 512 to 4096

In most cases, it is required that the overshoot does not exceed 5-10%, and in many practical problems it is not allowed at all. As we see from this example, the object of the second order in most cases can not be stabilized by a proportional regulator with a sufficiently small static error and without exceeding by overshooting of some pre-set not too large level.

Example 16. Consider under the same conditions a system with a third-order object (see Fig 4.8–4.9). As the proportional regulator coefficient *K* increases, the response time increases, the static error decreases, but the overshoot problem becomes more acute. As can be seen from Fig. 4.8, even with an object not prone to oscillation motions (as manifested in the real values of the roots of the polynomial of its denominator), control with a small overshoot can be achieved only with a small loop gain coefficient. In this example, this coefficient is equal to one, while the static error is 50%. Provision of a small static error (less than 0.1%) only at the expense of the proportional regulator is practically impossible with the majority of real objects of the third and higher order. Therefore, for these purposes, an integration channel is usefull, that is, a PI regulator.

In Fig. 4.10, the object previously considered is controlled using PI regulator. As is seen, a zero static error is easily achieved with overshoot of less than 20%. From the transient processes of Fig. 4.11 it is clear that the addition of the derivative link allows further reducing of the overshoot and eliminates the reverse overshooting for the prescribed value that took place in the previous example (see Fig. 4.10). In some variants of the values of the PID regulator coefficients, an additional increase in speed can be provided with full complete or almost complete elimination of overshoot, as Fig. 4.12 shows.

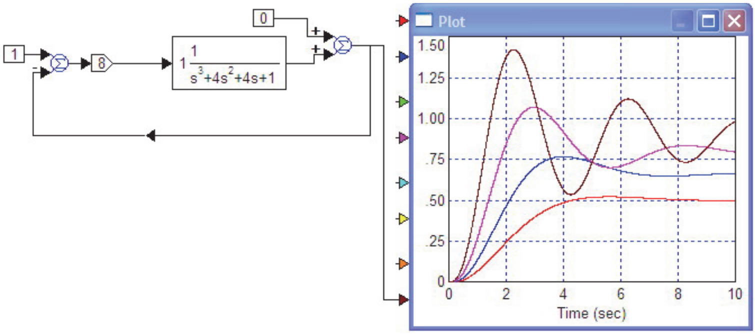


Fig. 4.8. Transient processes with a third-order object when the proportional regulator coefficient varies from 1 to 8

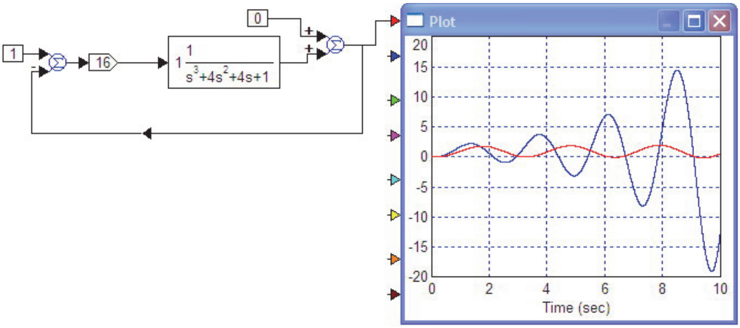


Fig. 4.9. Transient processes with a third-order object when the coefficient of the proportional regulator varies from 16 to 32

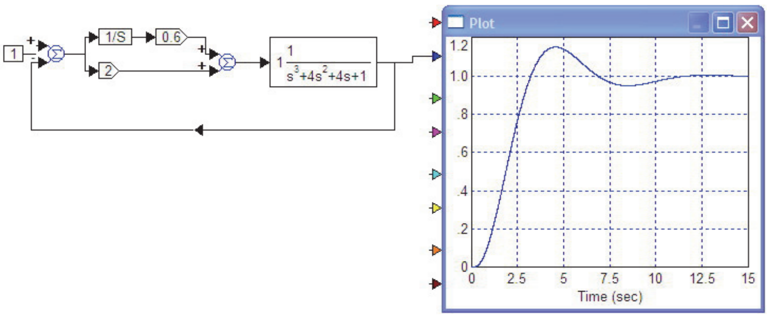


Fig. 4.10. Transient processes with a third-order object with PI regulator

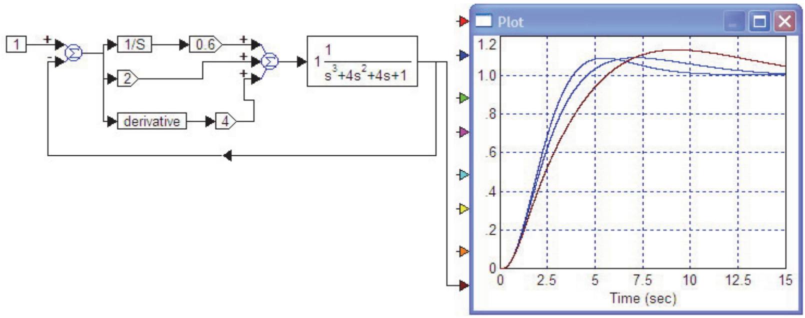


Fig. 4.11. Transient processes with a third-order object with a PID regulator

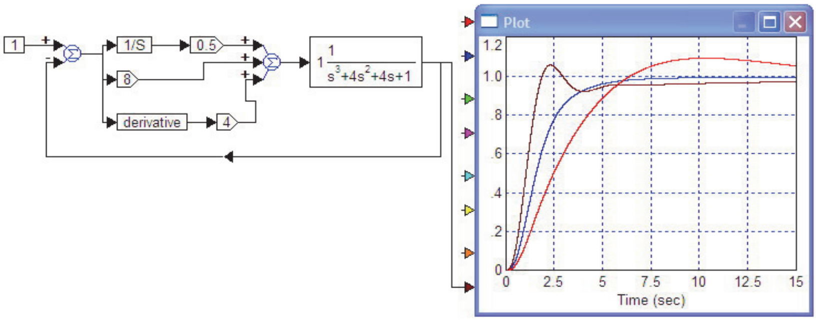


Fig. 4.12. Transient processes with a third-order object with a PID regulator with other settings of the coefficients

4.2. The differences between theoretical analysis, simulation and practical results

Consider a system with a transfer function of the loop which is described by a rational fraction. Usually, the order of object is the order of the polynomial in the denominator of the model of the object. The order of the regulator by analogy should be called the order of the denominator of its transfer function. Then the proportional regulator should be called the zero-order regulator, and the PI-regulator is the first-order regulator. In this terminology, the order of the numerator of the corresponding transfer function is not used, therefore this terminology is not sufficiently convenient.

For definiteness, we introduce the term “asymptotic order” (AO). We define it as the difference between the order of the denominator and the order of the numerator. This term is associated with the logarithmic amplitude-frequency characteristic (LAFC) slope in the high-frequency region; its value is equal to the slope multiplicity with the opposite sign. If AO is equal to unit, then the slope is minus one, and if AO is minus one, then the slope is plus one.

For example, the AO of an integrator, like the AO of an aperiodic link, is equal to one. In the high-frequency region, the slope of LAFC of these links is minus one (–20 *dB / dec*).

The AO of the oscillating link (the second-order link) is equal to two, the LAFC slope is minus two (–40 *dB / dec*).

For PID regulator AO is minus one. The slope of the LAFC in the high-frequency region is positive and equal to + 20 dB / dec.

Since, when connecting links, their transfer functions are multiplied (and LAFC are added), the AO of the corresponding transfer functions is also added. For example, when the second-order object and the PID regulator are connected in series, the resulting AO is one, that is, the LAFC of this connection in the high-frequency region is minus one.

From the theory of automatic control for linear systems it is known that a linear system without delay is stable if its LAFC crosses the axis of abscissas at an angle of –20 *dB / dec*. With sufficient length of this section (more than 12 *dB* in each direction), the system has sufficient stability margin.

From this it follows that the system, in the loop of which the total AO is equal to one, is stable with any gain.

Conclusion 9. A system consisting of first-order object with a proportional regulator is stable for any gain factors.

Corollary 1. To control the object of the first order, PID and PD regulators are redundant.

Conclusion 10. A system consisting of second-order object with a PID regulator is stable for any gain factors.

Corollary 2. To control the second-order object, the PID regulator is sufficient, double differentiation is not required.

The above conclusions and corollaries are valid only for the purely theoretical case. In practice, there is no first-order or second-order object. Such a simplified model is permissible only in the case when the error from such simplification is not fundamental and is negligible in the entire range of consideration of the behavior of this object.

For example, if the loop gain is limited for some reason, not related to finding of the optimal tuning, then the frequency range remains limited. Therefore, for the admissibility of such a simplification, it is sufficient that the simplified model coincides with the actual model of the object only in this limited frequency range.

A characteristic mistake of researchers is the violation of this rule. For example, when identifying an object, the adoption of a model of the first or second order is sufficiently justified, since in the considered frequency band this model was adequate. Further, this model is often used irrespective of bandwidth limitations, so as a result unreasonable calculations of the regulator are obtained.

It is sufficient to lay that under no circumstances any object has such properties that the system with this object and some regulator would be stable at an arbitrarily large gain. In any practical system, there is always a maximum permissible gain, above which it cannot be lifted, because otherwise the system would lose stability. From this alone, it already follows that there are never objects of the first or second order in nature.

There are some statements in literature, that the situation in theory is one, but in practice, it is quite different. Opposing the theory and practice is non-correct. Such statements are erroneous. If the theory does not coincide with practice, then a wrong theory is applied. Since the theory of automatic control is correct, in cases where theoretically calculated systems behave in practice differently from the theory, it is obvious that the error lies in the models of objects.

Conclusion 11. The model of the first-order object is always erroneous, since it is always insufficient for calculating the regulator.

Conclusion 12. The model of the second-order object is always erroneous, as it is insufficient for an accurate calculation of the PID regulator.

Indeed, with a PID regulator and a second-order object, the system is stable at any gain, which in practice cannot be.

In practice, there are no systems without transport delay. Accounting for even the smallest transport delay gives any theoretical model the properties of a real model: no delay model gives a system that is stable at any gain. A system with an object containing delay can always be brought to excitation when the gain is increased.

5. STATEMENT OF THE TASK OF OPTIMIZATION OF LOCKED STRUCTURES AND TOOLS FOR ITS RESOLVING

5.1. Demands to the locked loops of the model for simulation

Along with the requirement that all elements of the system should be physically realizable, as discussed above, it is necessary to select specific requirements for locked loops.

The first group of requirements is based on the possibility of modeling in software *VisSim*, the second group of requirements is based on the purpose of the control loop, which includes the regulator.

5.1.1. Demands for the possibility of simulation

In software VisSim, there is an additional requirement that in each loop, if to consider it in an open form, the degree of the denominator of the total transfer function was higher than the numerator, or there would be a pure delay in the loop. This requirement is based on the following. Each loop is locked through an adder, to the second input of which the signals come from the output of the same loop. There may be other inputs, which must also receive some signals. All signals arriving at the inputs of all adders must be determined by the beginning of the moment of calculating the output signal of this adder. If the loop contains inertial links, then at their outputs at the time of the simulation start the signal is zero because of their inertia. Therefore, all the initial data for calculating the output signals of the adders are available, and the system can be subjected to analysis by the simulation method in steps. The next steps by induction use those signals that were calculated earlier. If there is no inertia in the loop, the output signal of the loop, which is the input signal of the second input of the adder, must be known at the same moment when its input signal is calculated, which is the output signal of this adder. There is an unsolvable task for the program.

We can call as conditionally inertial element any element in which a time interval is not less than the integration step, which passes between the change in the signal at the input and the change that produces it in the output. This class of elements includes all filters, the degree of the denominator of which is greater than the degree of the numerator. In addition, this class includes delay links, in which the delay time is not less than the integration step. In addition, such elements are includes as sample and hold devices, registers and many other digital, analog and digital-analog devices. This class does not include derivative links, signal multipliers, gain factors, adders, logic elements without triggers, for example, elements “AND”, “OR”, “XOR”, and so on. We note that the differentiating element is formally one order of magnitude greater than the order of the denominator.

Accordingly, we can call as the inertial structure any connection of links, which collectively behaves like an inertial element. Accordingly, the combination of the inertial element and the non-inertial element gives the inertial structure except for the case with the derivative link. Sequential connection of the filter, the numerator of which is only one unit smaller than the denominator, with the differentiating link, as a result, is a non-inertial structure.

The above requirement can be formulated as follows: **every loop, if it is conventionally opened, must be inertial structure**.

This feature only at first glance may seem a shortcoming of the software. In fact, this limitation brings modeling closer to the real system, to the work of a real digital regulator. No loop in nature is non-inertial. Modeling of the non-inertial loop has no practical meaning.

In addition, the system indicates an error if, for example, a delay value is set in the delay link, which cannot be realized at the chosen integration step, for example, it is less than this step or not a multiple of this step.

5.1.2. Demands to the feedback loop from the point of view of the adequacy of the regulators to their tasks

The requirements for the system are formulated, as a rule, based on the technical characteristics that an object in the system should possess. As a rule, the following demands are required:

1. **Zero or negligible small static error**.

For example, it can be formulated mathematically:

$$\lim_{t \to \infty} y(t) = v(t) . \tag{5.1}$$

In addition, in the case of linear systems, this requirement can be expressed in terms of images:

$$Y(0) = V(0)\,. \qquad (5.2)$$

This requirement can also be expressed in the form of a requirement for the open-loop transfer function:

$$w < w_0 \Rightarrow |W_P(w)| \geq 10^J\,. \qquad (5.3)$$

Here J is a positive integer, for example, 10, w_0 is a small value of frequency, W_P is an open loop transfer function.

Fig. 5.1 shows a transient process with a steady error, which contradicts to the requirement 1.

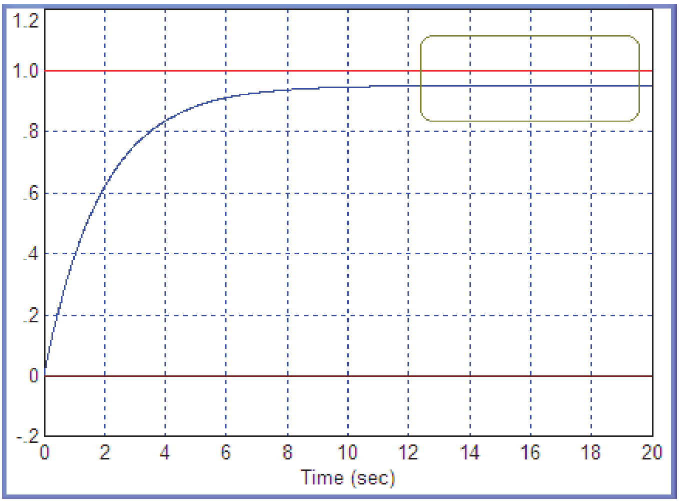


Fig. 5.1. Transient process with a fixed error

2. **A small overshooting that does not exceed a preset value, as a percentage of the value of the task jump that caused this overshoot**. In particular, there may be a requirement for no overshoot.

Fig. 5.2 shows a transient process with overshoot, which contradicts to requirement 2.

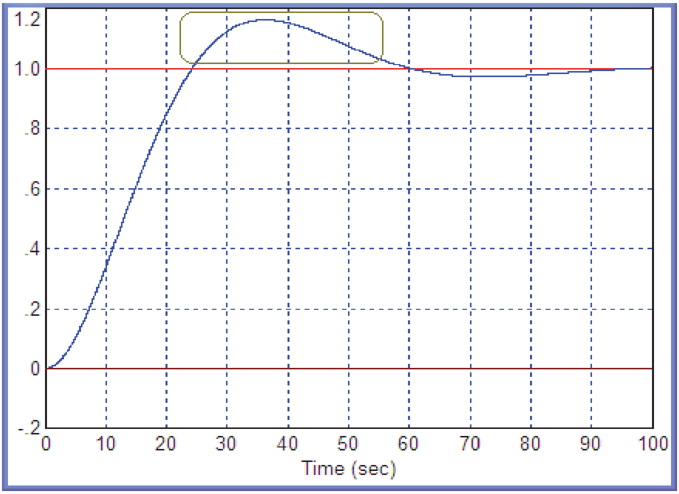


Fig. 5.2. Transient process with overshooting

3. **Short duration of the transient process** until error decreases to value less then certain small threshold. For example, the time to reach a relative error of 5% can be set:

$$t > t_{0.05} \Rightarrow | e(t) | < 0,05 \, v(t) \, . \tag{5.4}$$

4. **Absence of oscillations or their small number**, or the ratio of the amplitude of the next oscillation to the amplitude of the previous oscillation (damping index).

Fig. 5.3 shows a transient process with oscillations that contradict to requirement 4.

5. **No reverse overshooting**. The reverse overshoot is the output of the output signal in the direction opposite to the prescribed direction of this value.

Fig. 5.4 shows a transient process with a reverse overshoot, which contradict to requirement 2.

In addition, in a transient process, features that may worsen its attractiveness may occur, but often they are permissible, the presence of such features is not a reason to consider the system to be bad or inoperable. However, in some special cases, such features may be extremely undesirable.

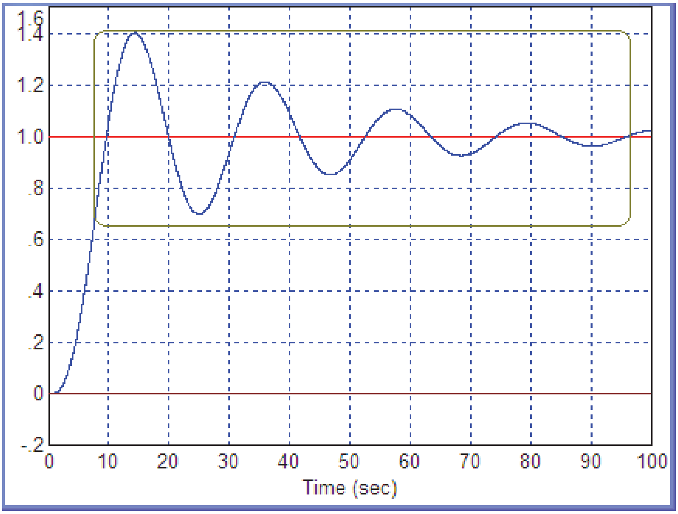


Fig. 5.3. Transient process with oscillations

6. **Not monotonic course of the transient process**.

Fig. 5.5 shows the transient process in a not monotonic fashion, which contradicts to the requirement 6.

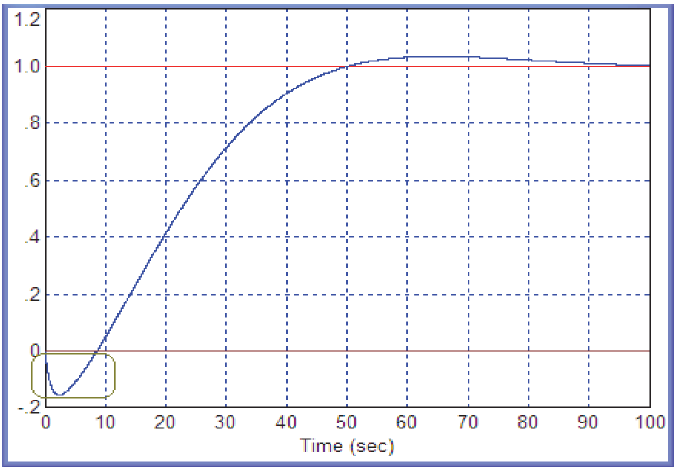


Fig. 5.4. Transient process with reverse overshoot

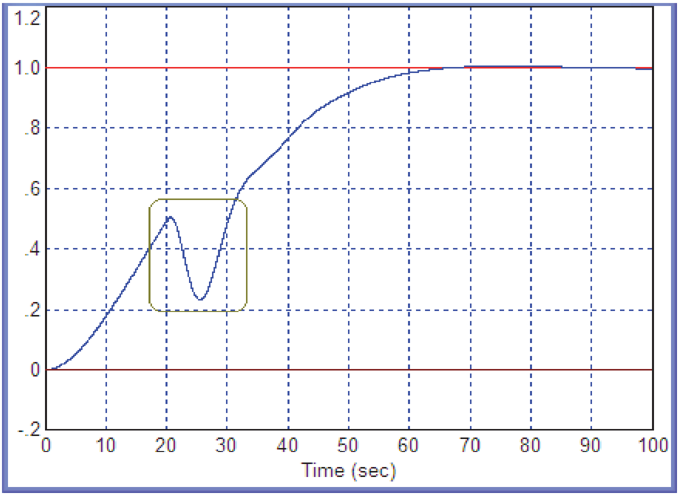


Fig. 5.5. Transient process with not monotonic motion

5.1.3. Demands to the cost function

The requirements to the regulator coefficients are formed as an objective function. The optimization procedure must find its extremum, that is, the maximum or minimum. If the procedure seeks for the maximum, the objective function is called the "Profit Function". If the procedure seeks for the minimum value, the objective function is called "Cost".

When using software *VisSim*, it is most convenient to use the cost function, that is, a positive-definite function that is calculated from the results of the simulation of the transient process, and which should take the minimum of all possible values for the computed arguments, in this case, with the found coefficients of the regulator.

If, according to technical requirements, the objective function is specified as a profit function, then it is easy to form a cost function from it, putting it equal to the value opposite to the profit function, i.e., profit with negative sign. That this function does not become negative, it can add a constant positive value, which is known to be greater than any current value of the profit. To treat lost profit as cost or prevented losses as profit is logical, this interpretation makes it easy to move from cost to profit and back, and, therefore, to ensure that the minimum of the cost function meets the best adjustment of the regulator.

At least, the following requirements are imposed on the cost function:

1. Cost function should depend on all the parameters that need to be optimized, directly or indirectly. It is not necessary that this dependence be expressed in an analytical relationship. It is enough that the relationship takes place.

2. Cost function must be a real nonnegative function for all values of all optimized parameters.

3. Cost function must correspond to the optimization objectives, namely: the better the optimization goals are achieved, the less is the cost function.

4. The dependence of the cost function on each parameter should be smooth and have a single minimum and the achievement of this minimum must correspond to the objectives of the control (the goals of regulator adjustment).

5. Cost function must not decrease indefinitely with an infinite growth (in absolute value) of at least one optimized parameter.

If cost function does not depend on even one of the coefficients, in the optimization process this coefficient will increase to infinity or decrease to minus infinity, the procedure will not end (the software will interrupt it), the optimization will not be successful.

If the dependence of the value function is not smooth, the optimization procedure will be unstable. This will be manifested either by the fact that at each attempt to continue the optimization from the achieved values it will lead to new values far from the initial values, or the result of optimization will depend on the starting conditions (or both of these phenomena will occur).

If there are several minima in the cost function, the optimization procedure can give different results, depending on the starting conditions. This defect can be corrected using the global optimization procedure, i.e. finding the global minimum of the cost function.

If the value function unboundedly decreases with the growth of at least one parameter, the optimization procedure will also not be completed, since it will be interrupted due to the fact that absolute value of this parameter will reach an unacceptably large value.

The procedure of global optimization can be obtained by applying the usual optimization procedure repeatedly from different starting values, storing the results (the received values of the cost function), and then choosing the result that corresponds to the minimum value of them.

5.1.3. Additional demands to the locked systems

In addition to the requirements discussed in the previous two sections, additional requirements may be imposed on locked systems. Examples of possible requirements are given below.

With respect to the object model, two possible options are possible: an **inaccurately specified object** and an **interval-defined object**.

In the first case, the parameters of the model are known, but with some error, for example, 1%. In the second case, it is known that at least one of the parameters of the object changes in a certain interval, the value of which is significant in comparison with its value, that is, for example, more than 10%.

To control an object with an inaccurately specified object, it is necessary to use **rough** systems. To control the object with interval parameters, it is necessary to use **robust** or **adaptive** system.

1. **Rough system** is a system that preserves efficiency and all design specifications with small changes in the parameters of the model of the object and the regulator. It is about changes at the level of no more than 2–3 %. This requirement is mandatory for all systems. The term "rough system" positively characterizes the system, and "not rough" means a negative property of the system, which consists in that there is at least one of the parameters, small changes of which lead to cardinal changes in the properties of the system (for example, the stable system becomes unstable, or overshoot in it increases twice or more substantially).

2. **Robust system**. The system is called robust if it is weakly sensitive to relatively large changes in the parameters of the object. In this case, one or more object parameters are specified in a certain interval. This change can not be attributed to the implementation error or the error of identification. Here we can talk about changing some coefficients by tens of percent or even many times, as well as changing the order of the model of the object. The regulator remains unchanged. It is this property, in contrast to "roughness", which we will later call robustness.

3. **Adaptive system**. In the case of changes in the model parameters that can not be attributed to inaccuracies in identification, it is required to provide the specified properties of the system, as in section 2, but it is permissible to change the parameters of the regulator, it is not required that the regulator remains unchanged. The structure of the system should contain a means for determining the parameters of the object, or the parameters of the system as a whole, as well as a means of influencing the regulator in order to change its coefficients or even the structure. Therefore, several types of adaptive systems can be distinguished.

A. **Adjustable (self-adjusting) regulators** (systems) are such ones in which the coefficients change.

B. **Switching regulators** (systems) are such ones in which the structure changes.

The property of “roughness” should be considered compulsory for each of the produced systems. Here we are talking about such a small change, which can be attributed to inaccuracy in the implementation of the regulator and inaccurate identification of the model of the object. Indeed, when manufacturing regulators using analogue techniques, as a rule, it is extremely difficult to ensure that the coefficients meet their calculated values with an error of less than 1%, and sometimes such an error can reach 5% or more. This is due to the accuracy of manufacturing resistors and capacitors and with the stability of their parameters. The use of digital technology allows some coefficients to be realized with greater accuracy, however, each digital device that is associated with the control object, as a rule, has analog nodes at the input and (or) at the output. Therefore, a regulator that is operable only with extremely precise implementation of all its coefficients and with extremely precise identification of the model of the object is useless. Consequently, this property must always be ensured. **Each object should be treated as an inaccurately specified object**.

The property “robustness” or “adaptability” can be considered as two different ways of solving the same problem, consisting in the system performing its functions, although the object model changes. Since robustness does not imply changes in the regulator, it is advisable to prefer such systems if their implementation is possible. Adaptive systems should be used when it is impossible to ensure the operability of the system without changing the regulator.

5.2. Structure for the regulator optimization

To optimize the regulator, it is required to apply a structure containing the system model and a number of auxiliary modules. The model of the system includes a regulator and an object, and other specific blocks can also be included.

In addition to the system model, the structure should contain:

1. Means of formation of test signals.
2. Means of indicating the results of optimization.
3. Means for calculating the value function.
4. Means of formation of initial values of parameters.

5. Optimization tool (performing the analysis of cost function, calculating new parameters, analyzing new costs and deciding on further steps - continuing the search or stopping it).

A possible structure for optimization is shown in *Fig*. 5.6.

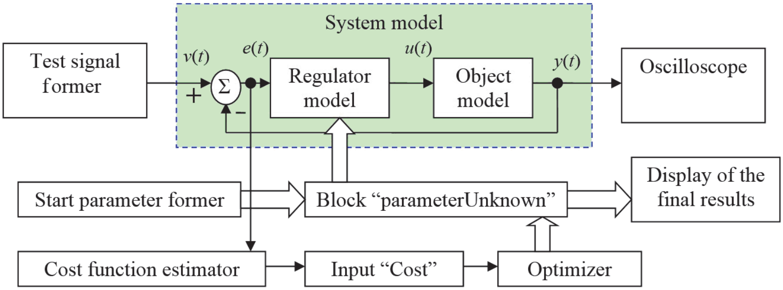


Fig. 5.6. A typical structure for optimizing of regulator of loocked system with feedback

This structure works as follows. The model of the object and the regulator model together form a model of the locked system. The means for generating test signals is a step jump generator, sine wave oscillator or something similar. For linear systems step is the most convenient and indicative kind of the input signal. The cost function evaluation mean calculates the value of the cost function based on the results of each test. The starting values are fed through the optimization unit to the regulator, specifying the values of its parameters at the very first test of the system from the series of model tests. The indicator means contains an oscilloscope showing the final transient process in the system, and displays the final values of the optimized parameters (regulator coefficients).

The optimization unit performs the following functions:

1. During the first test, it sets in the regulator those coefficients that are entered by the operator into the former of the start values, and also displays these values on the displays.

2. It started every next test.

3. Based on the results of the previous test, it takes in its input and analyzes the value of the calculated cost function in comparison with its previously obtained values.

4. It analyzes the stopping criteria.

5. If the stopping criteria are not met, then according to one of the selected algorithms, it calculates the new value of the set of regulator coeffi-

cients, and then returns to the repetition of all operations starting from step 2. If the stopping criteria are fulfilled, proceeds to step 6.

6. It selects the regulator coefficients corresponding to the lowest value of the cost function.

The stopping criterion can be one of the following:

a) the number of iterations has exceeded the specified value;

b) it is established that further implementation of this algorithm will not lead to a decrease in the value function by an amount exceeding the specified permissible optimization error;

c) an error occurred that makes it impossible to continue the procedure, for example, exceeding one of the parameters or one of the calculated values of the largest permissible value (for different versions of the program this is $10^{28} – 10^{30}$, etc.).

The cost function calculated in the structure implicitly depends on the regulator parameters and on the test actions generating the transient process, since the error function of the system $e(t)$ depends on these parameters. When solving the problem of numerical optimization, it is required to find such values of the regulator coefficients for which the cost function reaches a minimum value. Since the error in the system depends on the regulator coefficients, the entire cost function will depend on these coefficients. The solution of this problem is called the optimum, in our case the minimum. It is necessary to distinguish between local and global minima. Local minima are solutions in which small changes in any parameter cause an increase in the cost function; however, this does not apply to big changes in these parameters. The global minimum can be the only. This is the solution that ensures the least value of the cost function for all parameters from the range of their allowed values.

Various algorithms for finding the minimum can be applied. The correctness of the choice of the optimization method depends on the properties of the problem: the best method for one task may be the worst or completely inapplicable for another task.

5.3. Tools of the cost functions

In the simulation, various transient processes can be obtained depending on the values of the PID regulator coefficients. By comparing the various transient processes, developer can choose the preferred options and discard the unacceptable ones. Developer of the system always strives to ensure the greatest speed, the least overshoot and the smallest static error. If, however, the change in some coefficient causes improvement of one indicator, it simultaneously causes deterioration in another indicator of the quality of the transient processes. Hence it is difficult to choose the best option. In particular, the increase in the coefficient can simultaneously increase the speed and overshoot, and by the form of transients it may be not clear to developer which of the two variants of transient processes is most attractive. Therefore, a quality criterion is necessary that connects all other criteria into a single characteristic.

In some cases, developer can use two or more criteria, and even if there is no reason to choose one criterion from many, it will be sufficient to indicate that both transient processes satisfy the technical requirements for a locked system.

Still there is one problem, the solution of which can not be fulfilled without a single criterion for the quality of the system. This task is an **automatic iterative optimization of the regulator coefficients**.

We have already mentioned that the *VisSim* software tool can automatically optimize one or more parameters if there is a quality criterion. The criterion of quality can be any cost function that satisfies the requirements imposed on this criterion.

The cost function in general is written in the form of a functional:

$$\Psi(T) = \int_0^{\Theta} \psi(t)dt \, . \qquad (5.5)$$

Here Θ is the duration of the simulated transient process, under the integral there is a function that depends on time. As a rule, this function is associated with a transient process in the system during the development of a jump or another kind of change in the perturbation $h(t)$ or the prescription $v(t)$.

The most obvious, but not the best, option is as follows:

$$\Psi_1(T) = \int_0^{\Theta} e^2(t)dt \, . \qquad (5.6)$$

The following cost function is more effective in comparison with (5.6):

$$\Psi_2(T) = \int_0^{\Theta} |\, e(t) \,|\, dt \, . \qquad (5.7)$$

However, the relation (5.7) is lso not the best choice. The best value function can be computed in the case when the integral is the sum of several elementary functions with the corresponding weight coefficients. In general, this function can be written as follows:

$$\Psi(\Theta) = \int_{t=0}^{\Theta} w_q \sum_{q=1}^{Q} \psi_q dt . \tag{5.8}$$

Here, the value function is defined as the time integral of the weighted sum of the positive definite functions ψ_q, from the beginning of the transient process $t = 0$ to its end, when $t = \Theta$. The weight coefficients allow us to establish the ratio of the contributions of each of these functions.

One of the effective functions ψ_q for (5.8) is the error modulus $e(t)$ multiplied by the time t from the beginning of the transient process:

$$\psi_1(t) = | e(t) | t . \tag{5.9}$$

The use of such a function allows us to find the regulator, which most effectively reduces the error module, multiplied by the time. The expediency of reducing the error module of the justification does not require, and the multiplication of this quantity for a time is justified by the fact that the more time elapsed since the jump that caused the error, the better the remnants of this error should be suppressed. The initial value of the error is excluded from the objective function at all, because at this moment $t = 0$. The factor t plays the role of a weighting coefficient that continuously increases linearly. Time can also be applied to some positive degree, for example, t^2. This reinforces the requirement for rapid fading of the error and weakens the requirement for its magnitude at the very beginning of the transient process.

The disadvantage of the cost function based only on the term (5.9) in relation (5.8) is that often when using a regulator tuned by the optimization method with such cost function, oscillations in the transient process arise in the resulting system.

Several modifications of the objective function can be proposed to suppress oscillations. For example, the additional term may grow in the event that the overshoot exceeds a certain prescribed value.

For example, if it is required that the overshoot does not exceed 10%, then for a linear system this means that for a single step jump, the output signal, changing from zero to one, should never exceed 1.1. In this case, developer can form a "penalty" addition, which is equal to the positive part of the difference between the output value and a value equal to 1.1. If this difference is negative, then the penalty function is zero, if it is positive, then this value is equal to the value of the following function:

$$\psi_2(t) = \max\{0, x(t) - 1{,}1\}\,. \tag{5.10}$$

Here the function max $\{0, f\}$ is a limiter:

$$\max\{0, f\} = \begin{cases} 0, if\ f < 0 \\ f, if\ f \geq 0 \end{cases}. \tag{5.11}$$

This cost function is relevant only for the development of a unit stepped action, with other test signals it must be changed.

Another and more effective way to suppress oscillations in the transient process is the use of the error growth detector [31]:

$$\psi_3(t) = \max\left\{0, e(t)\frac{de(t)}{dt}\right\}. \tag{5.12}$$

The product of the error on its derivative must be negative for the best transient processes. In this case, that is, if the error and its derivative have different signs, the error value decreases during the process. The function (5.12) is zero, and its contribution to the value function (5.8) is also zero. This situation corresponds to the desired development of the process. If the error and its derivative have the same signs, the error in magnitude increases, the product of the error by its derivative is positive, and the function (5.12) is also positive. Then the value function (5.8) with the addition (5.12) under the integral increases as a result of integration of the positive function (5.12). The optimization procedure will find such regulator parameters that minimize the value with respect to the relation (5.8). Consequently, the procedure minimizes the parts of the transient process in which (5.12) is not zero.

This term (5.12) does not ensure the absence of transition regions on which the error increases, but it makes the contribution of such regions minimal, that is, minimizes their extent, and the value of (5.12) on them.

In the relation (5.6), the corresponding term has the form

$$\psi_4(t) = [e(t)]^2 . \tag{5.13}$$

It does not work well enough, but in some cases this cost function is justified. At the beginning of any process, the error is large. If the prescription $v(0) = 1$, then the initial error is $e(0) = 1$. During the rest of the process, the error modulus is much less than unity; therefore, the term (5.12) has an initial value that cannot be reduced.

In the general case, developer can use the error module to some extent, multiplied by the time to another integer power. Then the value functions will be the following:

$$\Psi(\Theta) = \int_0^{\Theta} | e(t) |^M \, t^N dt . \tag{5.14}$$

For $M = 1$, $N = 1$, we obtain a situation with the function from (5.9), for $M = 2$, $N = 0$ we obtain (5.6), and for $M = 1$, $N = 0$ we obtain (5.7). This is a positively defined function, that is, it cannot be negative for any values of its arguments. We investigated the dependence of the effectiveness of this cost function on the M and N degrees. It is shown that in the case of relatively simple linear objects this function works most effectively at $M = 1$, $N = 1$, and also at $M = 2$, $N = 3$ and in some other cases. In general, the growth of the exponent index M requires a stronger growth of the exponent index N. However, because of the possibility of using the composite value function (5.8), it is $M = 1$, $N = 1$, that is, relation (5.9), that can be recommended for most practical problems.

When creating a composite cost function, it is necessary to be guided by the principles of complementarities, competition and completeness. We shall call the integral of each of the terms under the integral in (5.8) a particular criterion, and the sum of these particular criteria will be equal to the value function as a whole.

The principle of complementarities

If any of the particular criteria works correctly, but not enough, then other particular criteria should complement its action. The contribution of each of them should influence the optimization result, which is ensured by the selection of weight coefficients.

The principle of competition

If a particular criterion acts on the result in such a way as to increase the regulator's coefficients, then this particular criterion will most effectively be supplemented by such a particular criterion that acts in the direction of decreasing these coefficients.

For example, terms that increase with a large overshoot in the system tend to cause the coefficient in the derivative link to increase it, but to decrease the coefficient in the proportional link. The criteria, depending on the dynamic error, act to increase the coefficient of the proportional link. A criterion that depends on the integral of the error, contributes to an increase in the coefficient of the integral link, but it may not be sufficient. For this purpose the cost function that contains the integral product of the error modulus for the time from the beginning of the transient process, as a term in the cost function (5.8).

Principle of completeness

If the regulator has an integral link, this may be because the system requires a zero static error. If the cost function does not include terms that increase dramatically when the static error is not zero, then the integral link coefficient will not be calculated reasonably enough as a result of the optimization procedure. If a second-order astatism is required from the system, it is not enough to introduce double integration into the regulator structure, it is also necessary to introduce into the cost function such a term that sharply increases if the system does not have a second-order astatism.

5.4. An example of the system quality analysis based on the cost function

Example 17. Fig. 5.7 gives an example of modeling the problem from Example 16 with automatic quality analysis based on the integral criterion from the control error $e(t)$.

The cost function is calculated from the ratio (5.14) with the values $M = 1$, $N = 1$, that is, based on (5.9):

$$\psi(T) = \int_{0}^{\Theta} t\left|e(t)\right|dt \,. \tag{5.15}$$

Since $e(t)$ depends on the coefficients of the proportional, integral and derivative links of the regulator P, I, D, the first condition is satisfied.

The purpose of regulation is to reduce the absolute value of the error $e(t)$ and the integral from it to a minimum. The smaller the value of $e(t)$ in absolute value, the smaller the value of function (5.15), therefore the second condition is also fulfilled. The factor t (time) under the integral is important. At the initial time, the error is inevitably equal to the value of the difference between the prescribed signal and the output, the action of negative feedback at this moment only begins to affect the behavior of the system. Therefore, the initial value of the error does not characterize the quality of the system. As the transient process develops, the quality of the system becomes more and more apparent. The faster the error decreases, the better the system is configured and calculated.

On the upper oscilloscope shown in Fig. 5.7, graphs of transient processes are presented, and the lower chart of this figure shows the changes in the cost function as these processes develop. In three cases from the four, the cost function continues to grow to the very end of the simulated transient process. This means that changing the simulation time would produces a change in the cost function, and, consequently, would change the result of the assessment of the transient process. To compare the results, it is necessary to made estimates only during the same simulation time and this time should be chosen so that by the end of it the growth of the cost function would ceases. This means that the simulation time must be chosen so that by the end of the simulation the error becomes zero or very close to zero in absolute value. Obviously, looking only at the lower chart, we would recognize that the best tuning of the system corresponds to the lowest value of the cost function, that is, just such that, after the tenth second, it practically stops growing. This corresponds to the graph on the upper chart, which is characterized by the absence of overshoot and the shortest duration of the transient process, that is, the output value reaches the prescribed single value most rapidly. Consequently, in this case the value function allows ranking the transient processes in adequate manner in terms of quality: the best value of the cost function corresponds to the best transient process. If the process were only considered for the first two seconds, we would prefer another cost function and another transient process. If the process were only considered for the first five seconds, then the cost function should prefer a process that has the highest initial speed, but at the same time has overshooting of about 5%, but by the type of process, perhaps we would prefer another adjustment of the regulator.

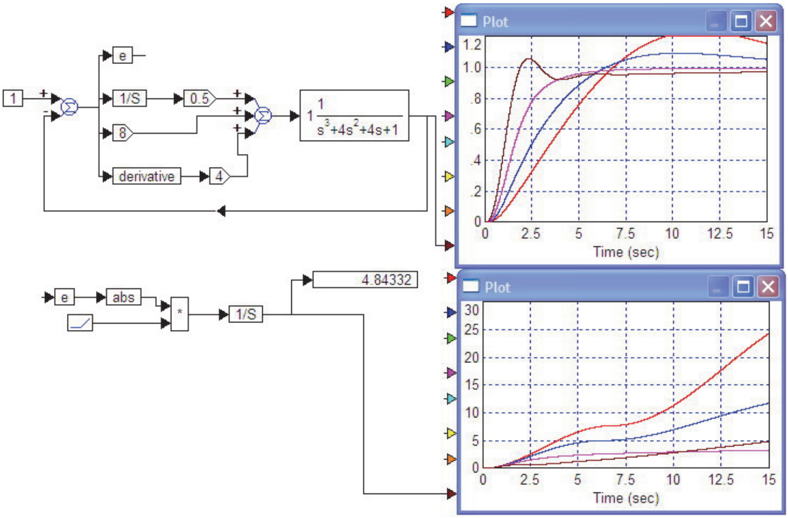


Fig. 5.7. Automatic analysis of the quality of transient processes with a third-order object with PID regulator using the integral quality criterion

This example shows how important is the correct choice of the cost function, along with the correct timing of the simulation of the transient process.

5.5. Grounds for choosing of weight coefficients in complex cost functions

The minimum of the sum of several functions, as a rule, does not correspond to a minimum of each of these functions. If there are several variants of objective functions, and there is a need to minimize all these functions together, then each of the terms must be in the sum with its own weighting factor. If the value would be unique, the coefficient would not play any role. When using two components in a cost function, at least one weight coefficient is necessary. To add different values it is necessary to bring them all to common units of measurement.

Choosing of the weights is also a procedure related to optimization. With different choices of these coefficients, the results of system optimization will be different, therefore, although each such result can be called "optimal", but only from the position of this choice of the criterion of optimality, that is, from the position of choosing weight functions. Thus, the optimization problem becomes more complex: first, it is required to "optimize" the optimization criterion, and second, based on this criterion, to optimize the regulator.

One way to choose weight coefficients is to find the general nature of the manifestation of these components, for example, "loss", or "price".

Another way to combine the criteria is to change the weighting coefficients based on the results of their application. That is, if a too large overshoot is observed in the transient process of the resulting system, then the weight coefficient of the term responsible for decreasing the overshoot must be increased. If, however, the error in the system does not decrease rapidly, then the coefficient of that term must increase, which increases sharply with the growth of the error, and therefore should provide a small error in the system. This method is effective, but it takes a lot of time and an informal analysis of the optimization result.

Example 18. Let there be a problem to optimize both time and cost of travel. Naturally, these costs can be saved only at the cost of each other, i.e. one can save time, spend more money, and save money, losing more time. The task of finding an acceptable compromise, in essence, boils down to the task of reducing the price of resources to a single price scale – or else one should estimate the time in money prices or money in hours, which, generally speaking, is the same.

Say, if the traveler earns 10 Euros per hour, we can assume that saving an hour is well worth it to pay at least half the price of hourly earnings. This gives the following natural criterion:

$$\psi_5 = \lambda_t \sum T_n + \sum \sigma_k \,. \qquad (5.16)$$

Here $\lambda_t = 5 \; Euros\,/\,hour$, T_k and σ_k are the time and money spent on travel of individual sections of the road.

In this case, one can use the expert method of determining the price. If it is difficult to name the price of a certain work, then it is usually sufficient simply choosing of two limiting prices, namely: surely low and surely overpriced.

Example 19. It is necessary to determine how much one hour of your non-working time is. Let us say you would have given 4 Euros without a doubt. For saving this time, but under no circumstances would they give for it 100 Euros. Therefore, the true price for you is somewhere in the middle. "Mean" can be used as the arithmetic mean and as geometric mean with an equal degree of validity, if there is no special reason for preferring one or another method of averaging (geometric mean, this is the average of the logarithm). In this example, the arithmetic mean gives 52 Euros, and the geometric mean gives 20 Euros. If we continue to find it difficult to choose a particular answer, we can calculate the average between these two results. The average of the arithmetic between these results is 36 Euros. The average geometric is 32.25 rubles. These results are almost close. We can say that the cost of one hour under the initial conditions is 34 Euros.

If the extreme estimates differ by orders of magnitude, for example, 10 and 100000, it is better to use geometric mean, if they are comparable in order of magnitude, for example, 10 and 60, then the arithmetic mean can be used.

6. NUMERICAL OPTIMIZATION OF THE LOCKED STRUCTURES

6.1. Procedure for the automatic optimization of regulators

To run optimization in *VisSim*, a developer should do the following things:

1. Setting of one or more optimized parameters using the "parameter unknown" blocks.
2. Setting the initial (starting) values to each of these parameters, giving a constant to the input of these blocks.
3. Ensuring of the use of these parameters as regulator coefficients to be optimized. For this purpose, it is useful to give them the names of functional variables (for example, *P, I* and *D* are the coefficients of the proportional, integral and derivative links). Assigned names should be used to call these functions as appropriate gains. A developer can use multiplier blocks or directly blocks of gain factors.
4. Ensuring of calculation of the cost function and sending of the result to the “cost” block, which should be the only one in the project. For this goal, a calculator is formed, the output of which is connected to the input of this block.
5. Providing of an indication of the optimization result, for which it is advisable to use the value indicator block at the output of each of the “parameter unknown” blocks.

In addition, it is possible, but not necessary, to reflect the value of the resulting transient process on the oscilloscope block and the value of the cost function on such a unit and (or) on the numeric value indicator unit.

Example 20. Fig. 6.1 shows an example of a PID regulator optimization project for an object from Example 16. The criterion (5.15) is used. The resulting coefficients should be rounded to values containing no more than three or four valid digits for each parameter. In this example, the starting values of the coefficients of the integral, proportional and derivative links were, respectively; $k_p = 1$; $k_i = 0$; $k_d = 0$, and after optimization we have obtained $k_p = 0.584$; $k_i = 3.02$; $k_d = 1.9$.

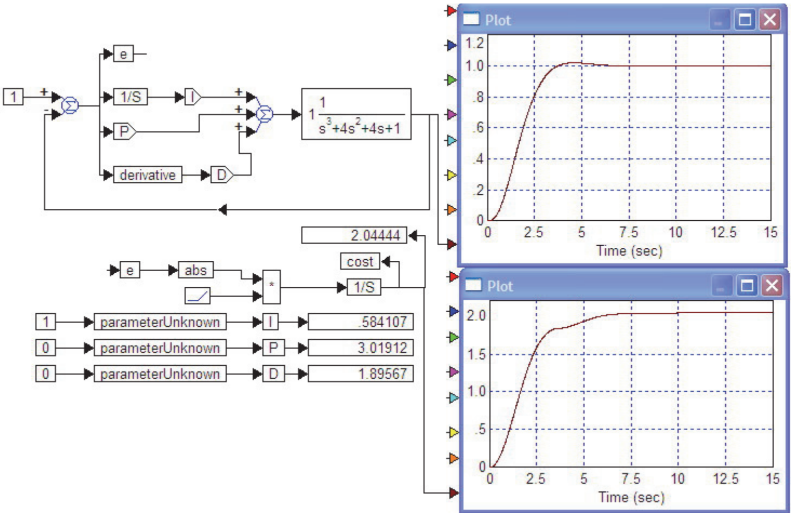


Fig. 6.1. The project of automatic optimization of the PID regulator for the object from Example 16 by the criterion (5.15)

After rounding the obtained coefficients, it is necessary to verify the admissibility of such rounding and the preservation of the achieved type of transient processes (minor changes are allowed) and approximate preservation of the value of the cost function. To do this, it is necessary to fill the blocks of preset values by "copy-paste" method with the received parameters values and round these parameters. In "optimization" menu one should check the "optimize" window, select the "remember graphs" checkbox in the oscilloscope configuration menu and start the simulation again.

This method of verification is necessary to exclude a non-rough solution, i.e., a solution that ensures a given quality only if the values of the parameters are extremely accurately matched to their values.

The presence of a non-rough solution can arise in the case, for example, when an object contains an unstable oscillatory link, for the stabilization of which mathematical modeling suggests the introduction of a blocking filter tuned to the same frequency in the regulator. Such a decision has no practical value, since its implementation is impossible, and also because a sufficiently accurate identification of such object is also impossible. An object with such properties is most likely (not exactly) with the model used, but the implementation of the regulator, most likely does not coincide with the calculated regulator.

Example 21. Fig. 6.1 shows an example of a PID regulator optimization project for the same object from Example 16 using the criterion (5.6). As it can be seen, the result of the decision is not so attractive. Therefore, cost function (5.6) can be considered less successful for this problem. However, we note that our analysis of the transient process is carried out visually according to the type of the transient graph. If the use of this criterion was dictated by the technical conditions of the operation of the object (and the system based on it), for example, it is the square of the error that characterizes the amount of produced defective products or other adequately represented types of losses from insufficient control accuracy, then this criterion should be used. In this case, not our subjective assessment of the transient process deserves more confidence, but the result of calculating the value function (5.6).

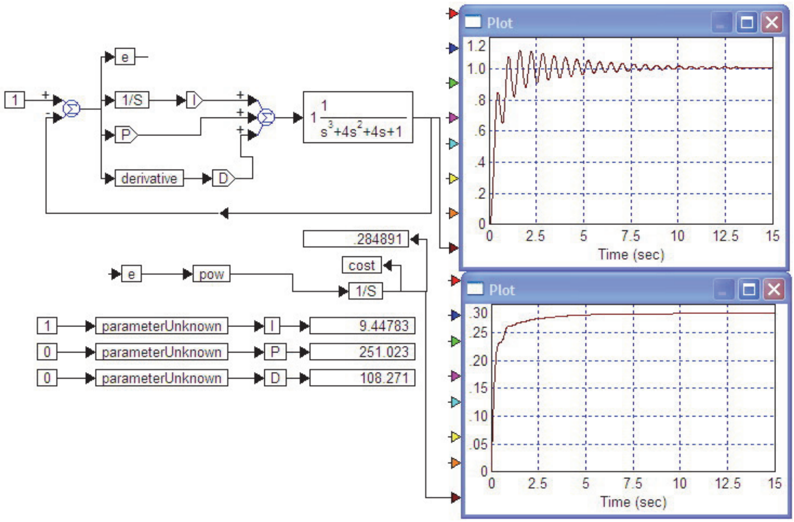


Fig. 6.2. The project of automatic optimization of the PID regulator for the object from Example 16 by the criterion (5.6)

Example 22. Fig. 6.3 shows an example of a PID regulator optimization project for the same object using cost function (5.7).

As we can see, the criteria (5.15) and (5.7) provide sufficiently attractive transient processes with the selected object from Example 16.

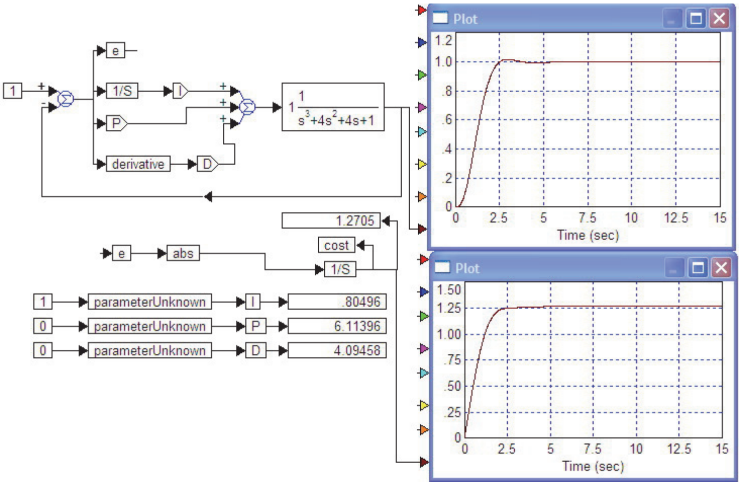


Fig. 6.3. The project of automatic optimization of PID regulator for the object from Example 16 by the criterion (5.7)

Let us note some more important principles of application of automatic optimization.

1. Not only the **step** of sampling in time, but also the **simulation time (duration) should be chosen reasonably**. We should strive to ensure that the transient process ends in approximately 70–80% of the simulation time. If the transient process is not yet finished, the optimization results should be checked. If, however, the transient process ends too quickly with respect to the simulation time, hence, the cost function includes the result of the signal processing, which is weakly related to the quality of the transient process.

2. It is advisable to strive to ensure that the cost function at the end of the transient process would reach a stable level and completed its growth by some steady-state value. If this is not the case, then, the choice of simulation duration makes a significant contribution to the cost function and, therefore, to the optimization result, too. The growth of the objective function at the end of the simulation time may also indicate some errors or incorrectness of using the optimization mode, for example, inadequate consideration of factors such as noise or the ADC sampling error that performs the measurement conversion of the output signal.

3. If optimization is carried out using a real object or using the most detailed model, up to the ADC or noise, it should be noted that **the control error cannot be strictly zero in practice**. In this case, it is necessary to use

the maximum permissible error as the boundary value. In any real system, one can specify a sufficiently small value of the error, which can be considered **negligibly small**, that is, the absolute error value falling into the interval between zero and this value can be equated to zero error value. In this case, such a situation can be considered an achievement by error of zero level. It is necessary to use this in the objective function, that is, **to introduce non-linearity of the "dead zone" type**. Then the integral of such an error will be zero, which corresponds to the natural acceptance of such an error inessential for the operation of the system.

4. As a result of optimization, a **structurally unstable system can result**. It should be ensured that the resulting system is sufficiently **rough** so that it can be used in practice. To do this, it is sufficient to round off the coefficients obtained, leaving only 2–3 significant digits in each of them, and repeat the simulation. The optimization procedure should not be repeated. The resulting transient processes should not differ too much from the processes obtained with the exact values of the regulator parameters calculated as a result of the optimization procedure. If this is not the case, then the system is not rough, and its results are not reliable, that is, the resulting regulator cannot provide the required quality of functioning of a real system.

Example 23. Consider the problem of optimization of a PID regulator for an object with the following transfer function:

$$W(s) = \frac{\exp(-10s)}{(20s+1)(s^2+2s+1)}. \tag{6.1}$$

The corresponding structure of the object is shown in Fig. 6.4. The structure of PID regulator is shown in Fig. 6.5. The structure for optimizing of PID regulator as a whole is shown in Fig. 6.6. In this case, the object and the PID regulator are assembled into composite blocks for more understandable picture.

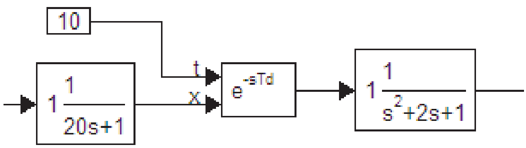


Fig. 6.4. The structure of the object of automatic optimization of PID regulator from Example 23

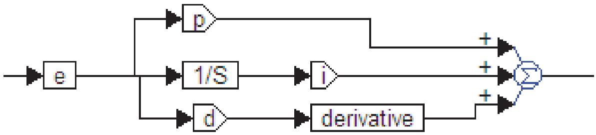


Fig. 6.5. Structure of the PID regulator (for all cases)

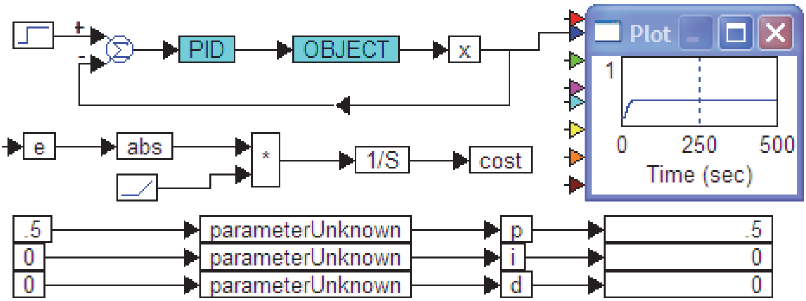


Fig. 6.6. The structure of the system for automatically optimizing the PID regulator from Example 23

When the optimization procedure is started, the values of the required PID regulator coefficients are obtained in the windows of the optimization unit's display, as shown in Fig. 6.7. The resulting transient process is shown in Fig. 6.8. The transient process ends in about 70 *s*. The overshoot is negligible (about 2%) and the static error is zero (which is natural when using the PID regulator).

The next step is to check the result for roughness. For this goal, we round off the results. We set $k_p = 1.5$; $k_i = 0.052$; $k_d = 6.5$. The resulting transient process almost completely coincides with the process shown in Fig. 6.8, from which it follows that the system is rather rough.

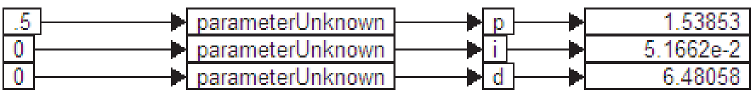


Fig. 6.7. The result of optimizing the PID regulator from Example 23

It is also possible to demonstrate (but not prove) the optimality of the obtained solution. To do this, it is sufficient to add an indicator of the value of the cost function to the structure, after which to change the individual coefficients in the direction of increasing and decreasing. If the system is optimal, then any change in any coefficient will give an increase in the cost function. With the obtained coefficients, the value of the cost function is 231.5.

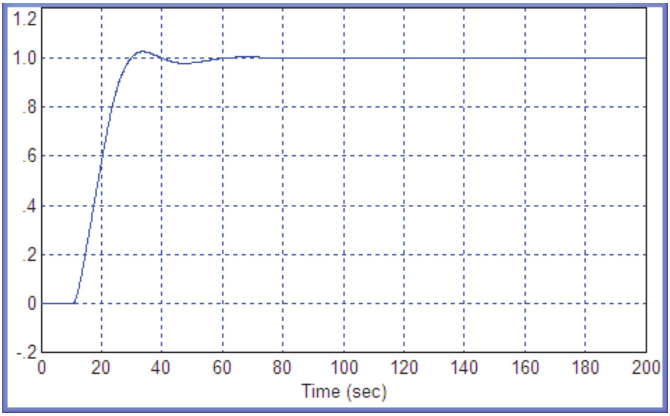


Fig. 6.8. The transient process resulting from the optimization of the PID regulator from Example 23

We can set the increment to the proportional coefficient by 0.2. When this coefficient is increased by this value, the value function becomes 292.0, and with a decrease it is equal to 371.8. The graphs of the corresponding transient processes are shown in Fig. 6.9.

Similarly, we set the increment to the coefficient of the integrating path by a value of 0.04. With an increase in this coefficient, the value function becomes 285.5, and with a decrease of 277.5. The corresponding transient processes are shown in Fig. 6.10.

Let us change the coefficient of the derivative path by 0.4. With the increase of this coefficient, the value function becomes 244.3, and with a decrease it is equal to 222.7. The corresponding transient processes are shown in Fig. 6.11.

The latter actions, seemingly, refute the assertion that the obtained values of the PID regulator coefficients are optimal. However, the conclusions should not be rushed. The point is that we rounded out the optimization results. If we take more precise results, namely, $k_p = 1.538$; $k_i = 0.0516$; $k_d = 6.4$, then the value function takes the value 217.0.

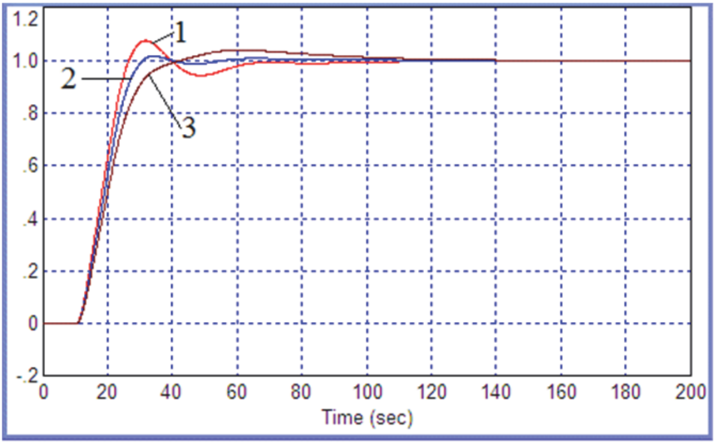


Fig. 6.9. Transients from Example 23 with proportional link coefficient changes: 1 – increase, 3 – decrease, 2 – initial value

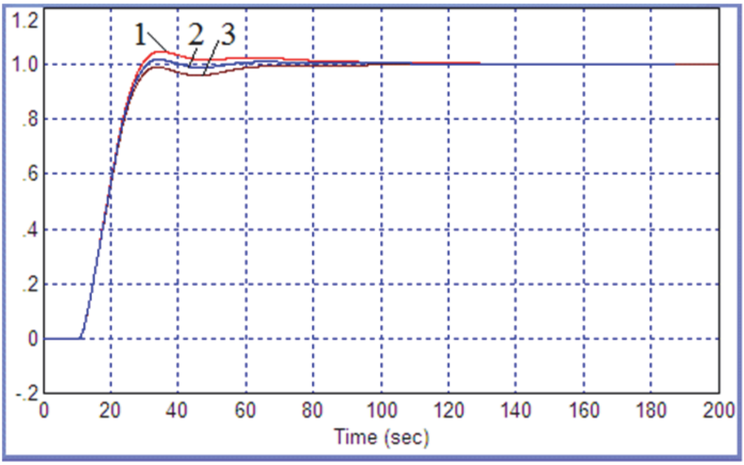


Fig. 6.10. Transients from Example 23 with changes in the coefficient of the integrating link: 1 – increase, 3 – decrease, 2 – initial value

Therefore, it nevertheless should be recognized that the tests carried out have confirmed the optimality of the result obtained (although this acknowledgment refers to the result with an accuracy of four significant digits). The changes in the graphs of transient processes during rounding are not significant, so the rounded value of the regulator parameters can be considered optimal in some approximation.

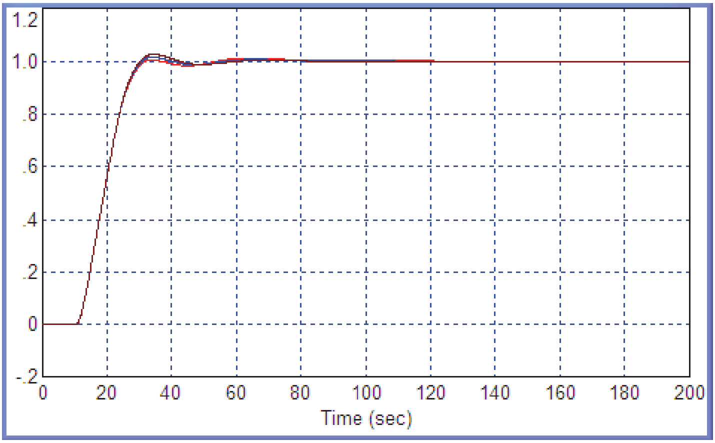


Fig. 6.11. Transient processes from Example 23 with changes in the coefficient of the derivation link: all lines merge

6.2. Fixing of the coefficients

If in the use of an insufficiently high version of the software *VisSim*, or in the optimization there are of too many regulator parameters, or in other cases which complicate the optimization of a regulator, it may turn out that the program is not able to find all the required coefficients. This can be expressed in the fact that the resulting transient process is not satisfactory, or the system does not terminate the procedure in spite of the fact that the number of iterations made has clearly exceeded the preset maximum value.

One of the methods of simplifying the problem is that some coefficients can be fixed by removing the extra blocks "parameter unknown". After finding a part of the sought coefficients, a developer can change the choice of which factors are fixed and which ones change during the optimization procedure.

Also it is possible to fix the correlation between some coefficients, which also allows reducing the number of required parameters. For example, if certain symmetry is observed in the structure of a multichannel object, we can expect an analogous symmetry in the multichannel regulator, which allows us to use, for example, equal regulators in symmetric channels. This also reduces the number of search parameters in the optimization procedure.

6.3. Optimization of the ensemble of systems for the robust control

In the simulation, several parallel systems can be created. Two cases are possible: a) parallel modeling of systems with the same objects and different regulators; b) parallel modeling of different objects with the same regulators.

In case (a), this can speed up the process of finding the extremum and make it more visible. For example, if we do not know whether this regulator parameter should be increased or decreased, we can try to do both in sequence, and we can implement both in parallel. In the case of finding the optimal setting with a real object, a developer often has to perform a trial deviation of the tunable parameter. This deviation allows determining of the direction of further changes to this parameter. In the simulation, a developer is not limited to one object, he can simulate two or even more objects in parallel. This eliminates the need for trial deviation.

Case (b) allows the calculation of robust regulators, i.e., such regulators, that provide a satisfactory solution of the problem not only for fixed parameters of the object, but when they change in a certain small (previously known) interval.

Indeed, if a regulator is necessary which successfully works with the object under the conditions of changing its parameters in a certain interval, then developer can simulate several systems in parallel with identical regulators, but with different objects. At the same time, different settings of object parameters correspond to extreme and (or) most undesirable combinations of object parameters. The regulator, which will give the best results for all objects from this list at the same time, will be robust, in the sense that it will be successful in a predetermined range of object parameter changes. The quality criterion (cost function) must take into account, equally (or, possibly, with weight coefficients) the quality of each of the systems being modeled.

Example 24. Consider the problem of optimizing a PID regulator for an object from Example 23, provided that the time constant of the delay link can increase to a value of 15, and the time constant of the first filter can increase to a value of 60.

If we use the result obtained in Example 23, then at the highest values of the indicated time constants the transient process will be characterized by a 30% overshoot, as Fig. 6.12 shows.

Optimization of the ensemble gives the transient processes shown in Fig. 6.13. It is seen that the process for smaller values of the time constants (line 1) converges too long, since the error does not become zero until the end of the graph (that is, until the time 200 *s*). We can enter a weighting factor equal to 1.5 into the error from this system when calculating the cost function. The corresponding structure is shown in Fig. 6.14. The process is shown in Fig. 6.15. It is clear that as a result, the both processes reach the prescribed value (the error becomes zero) in a time of about 160 *s*. Thus, the optimization of ensemble of systems allows ensuring of the achievement of a compromise in the quality of transient processes at the boundary values of the model parameters of the object.

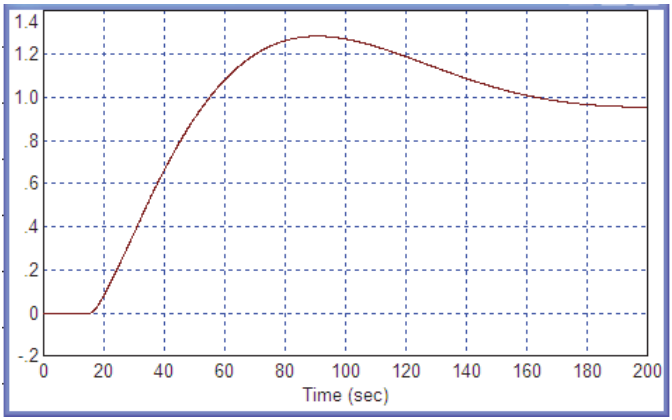


Fig. 6.12. The transient process in the system of Example 23, obtained in a system with an object with maximum time constants when using a regulator obtained by optimization with the minimum values of these time constants

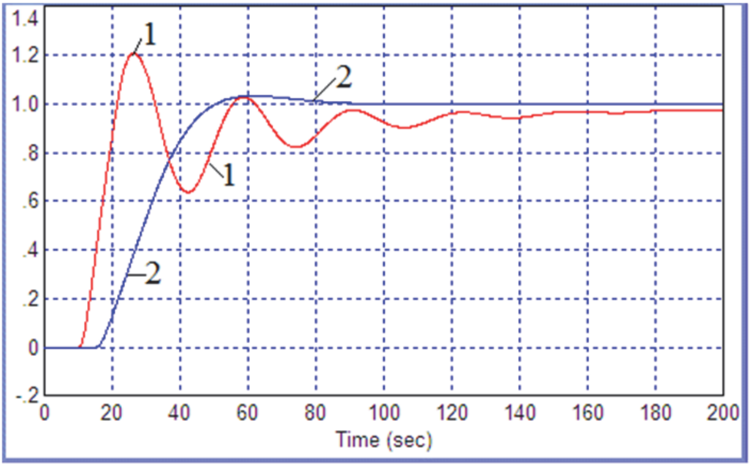


Fig. 6.13. Transient processes in the system from Example 23, boundary values of time constants (line 2 is the largest values, line 1is the smallest values); the regulator is calculated by the method of optimizing of it for the ensemble of the systems

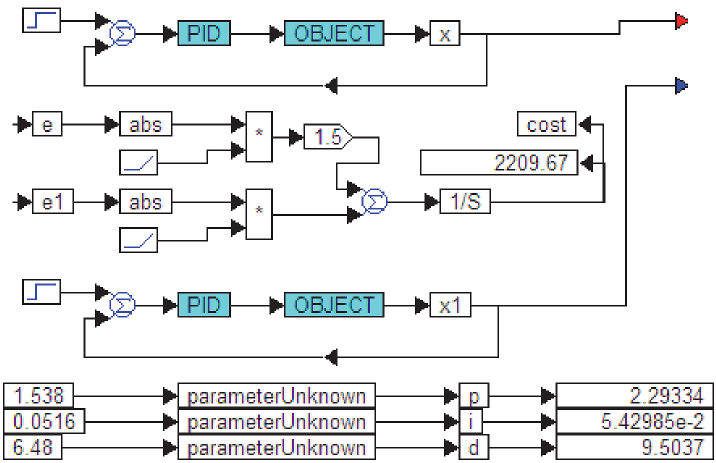


Fig. 6.14. Structural diagram for optimizing the regulator for an ensemble of systems using a weighting factor in the cost function

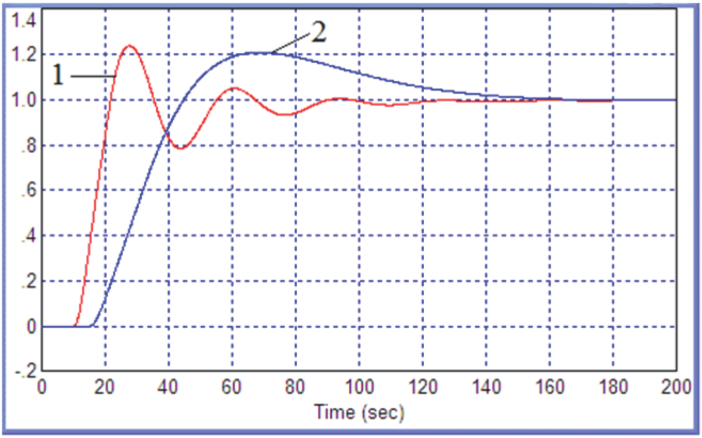


Fig. 6.15. Transient processes in the system according to the structure of *Fig*. 6.14: line 1 is output of the object with smaller parameters, line 2 is output of the object with large parameters of the model

6.4. The validity of the model for optimizing of the regulator

The need for the adequacy of the model used to its actual value is discussed earlier. It is required that the regulator, which is optimal for the used model, would remains optimal for the real object. Every model is approximate; no model corresponds to the object completely. Therefore, the question of sufficiency of approximation of the model to the real object is relevant. We need a formal criterion for the admissibility of neglecting certain parameters of the model, and so on.

As noted above, the object of control is actually considered in a limited frequency band. With this stipulation, it is permissible to describe objects by first- or second-order models, as well as models that do not contain inertial links, and even by derivative links, i.e., elements, the order of the numerator of transfer functions is higher than the order of the denominator. But in this case, when solving optimization problems, the object and its model can be identified only in the specified limited frequency region. Otherwise, there is an apparent contradiction between theory and practice, namely: the optimal regulators in theory prove to be non-optimal in practice and vice versa.

In order to avoid the contradiction between theory and practice, it is necessary to artificially limit the frequency band in the model in accordance with how it is naturally limited in a real object, or (more correctly) to refine the model in the high-frequency region to such a sufficient degree that it would ensure a coincidence of the theory and practice.

In Section 4.2, the term "asymptotic order" (AO) of the open-loop transfer function is introduced. With respect to the loop with AO of the first order, the theory at first glance differs from practice.

In theory, it is impossible to optimally adjust the loop with a first-order AO. **In practice**, for any system, optimal tuning is possible. This contradiction is resolved as follows: **in practice, there are no first-order AO loops**.

The impossibility of optimization does not mean that the object can not be controlled. It may be that just in case the object is too simple to control, the regulator can not be optimized. Optimization involves a compromise, that is, the choice of such coefficients, which are not too large and not too small. For example, the gain in the regulator should not be too small, so that the duration of the process and the error are not too large, but this gain should be not too large, so that overshooting would be not too large. If, however, the increase in the gain makes it possible to reduce the error and the duration of the transient process, but does not lead to a violation of stability and does not even lead to the appearance of overshoot, then a search for a compromise is not required, and it is impossible. In the theory of such a system, an increase in amplification gives only a positive effect without any negative side effect. Therefore, the impossibility of optimization in theory does not indicate the fundamental complexity of solving the problem of controlling an object. It only points out the complexity of solving this problem by the optimization method. But this complexity is not fatal: it is sufficient to implement some additional constraint that will cause the optimization procedure to stop at some stage, and as a result, a regulator adjustment will be obtained that will be satisfactory. It is required that the reason for stopping the procedure introduced artificially with the expansion of the speed of the system into the high-frequency region should act no later than the actual reason for the impossibility of increasing the gain in the real system.

Thus, when using a procedure that does not end itself, developer should automatically stop earlier than the procedure will lead to the area of inadequate solutions. This principle can be formulated as follows: an artificial obstacle must act before the catastrophe occurs. The catastrophe here means an erroneous system setup in which a system with an erroneous model is stable, and a system with an actual model is unstable.

On this basis, we can propose the use in the optimization procedure of a model with an artificially imposed restriction on its speed.

So, let the optimization problem of a PI regulator or proportional regulator for a first-order object is not correct, since there is no reason for the formal procedure to stop the gain increasing. Similarly, the problem of optimizing the PID regulator for objects of the first and second order is incorrect. This means that whatever setting we have not found, there is always the option to specify a different setting, which will be better for the quality of achieved control. The introduction of the artificial stopping mechanism of this procedure makes it stable; the task with this modification becomes correct.

An artificial obstacle for this procedure can be done, at least, in one of the following two ways:

1. Restriction on the growth of one or several gains.
2. Restriction on the speed of the object.

In the first case, for example, a value can be introduced into the cost function, which is equal to zero if the product of all the coefficients is less than a certain prescribed large value.

In the second case, for example, a link with a limited speed can be introduced into the model of the object.

The second option seems more natural, since it is more real situation in practice.

Therefore, in practice, for any object without exception, there is always the best setting (regardless of the possibility of finding it). This is due to the fact that in practice, there are no ideal objects of the first or the second order, or any other finite order in an unlimited range of frequencies and without consideration of the limitations of the input effect. It is these limitations that often lead to the fact that after achieving some acceptable quality, no methods can improve the result, and this is why a solution that cannot be improved is called optimal. The reasons for the existence of an optimal solution can be called the reasons for the limited possible speed and quality of the transient process. They may optionally be associated with a limited input impact or a limited bandwidth: another reason may be a transport delay in the real-world model.

In practice, **there are all these reasons and some others**, but **not all of them have a determining effect on the optimal setting**. The one that starts to affect them at lower frequencies (earlier) is the main reason for the model's inadequacy, and it must be taken into account in the refined model. Which of the reasons is most important, it depends on the ratio of the corresponding values in the true mathematical model of the object.

Such a cause may be a delay link, or additional unaccounted minimum-phase links in the object model (i.e. filters), or dynamic non-linearity, that is, a limited rate of increase in the output signal, and so on.

6.5. Forced limitation of the regulator coefficients

In the previous section, it is said that an incorrect task can be made correct if the regulator's values are forced to be limited in magnitude.

Example 25. Consider the problem of optimizing a PID regulator for an object with the following transfer function:

$$W(s) = \frac{1}{100s + 1}. \tag{6.1}$$

The structure for optimization and the result (the obtained coefficients) are shown in Fig. 6.16. The stopping of the optimization procedure was because an integration step of 0.1 *s* was selected. The resulting transient process is practically discrete; it ends after a small number of simulation steps, as shown in Fig. 6.17. This result cannot be satisfied.

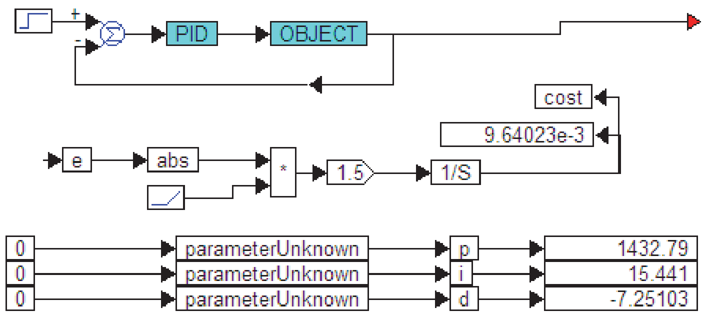


Fig. 6.16. The block diagram for optimizing the regulator in Example 25

We reduce the integration step by a factor of 10, that is, we select 0.01 *s*. We obtain new regulator coefficients, equal to: $k_p = 10000.5$; $k_i = 100$;

$k_d = 4.137 \cdot 10^{-11}$. The resulting transient process is shown in Fig. 6.18. It is again practically discrete. If we further reduce the integration step, we will obtain even larger coefficients with even greater speed, and again the procedure will stop when the process is actually discrete.

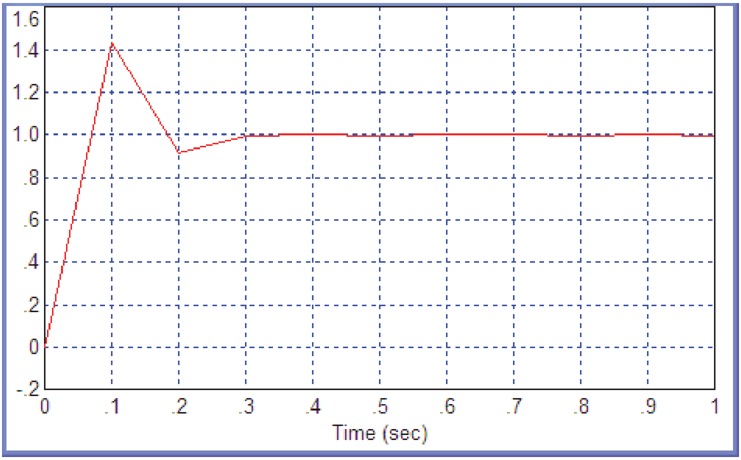


Fig. 6.17. The transient process in the system optimized according to the structure of Fig. 6.16 with integration step of 0.1 *s*

Let us introduce a restriction on the gain of the proportional gain loop, for example, at the level of 1000. For this goal, we introduce into the cost function the integrated signal former of value k_p, the non-linear element "dead zone" with the band width value 2000 and the absolute value calculating unit. We connect the output of this block through the adder to the integrator. The new structure of the system model is shown in Fig. 6.19.

As a result of optimization, a regulator with the following coefficients was obtained: $k_p = 1000$; $k_i = 23.377$; $k_d = -57.21$. The resulting transient process is shown in *Fig*. 6.20. This method succeeded in making the incorrect task correct, if we now reduce the integration step, the result will be the same or very close. Indeed, if the integration step is reduced by half, the result is almost the same, namely: $k_p = 1000$; $k_i = 33.02$; $k_d = -69.71$, the type of the transient process almost does not change, it only shortens by 20%.

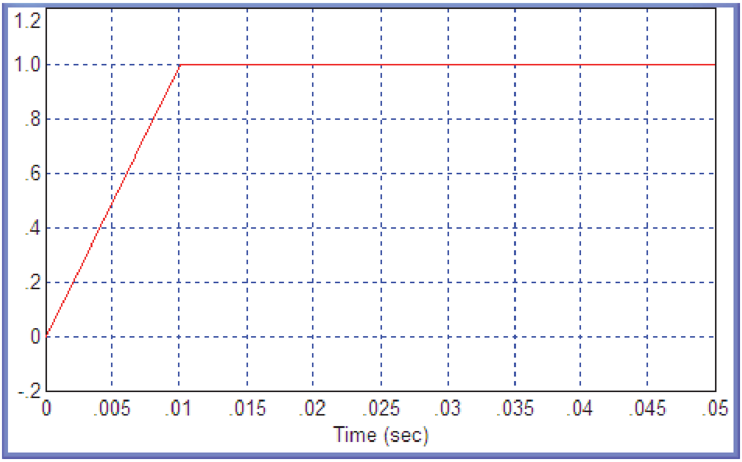


Fig. 6.18. The transient process in the system optimized according to the structure of *Fig*. 6.16 with integration step of 0.01 *s*

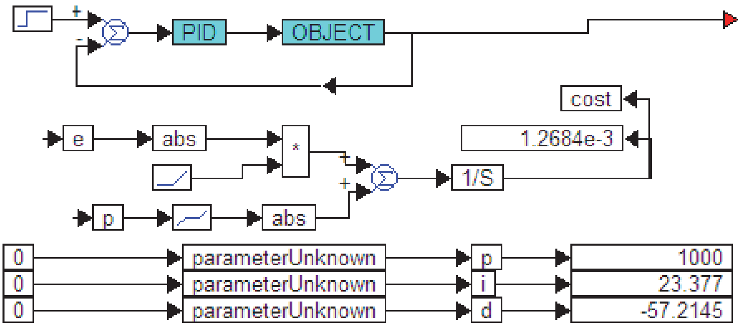


Fig. 6.19. Modified block diagram for optimizing the regulator in Example 25 with limiting the proportional gain factor of the regulator

Therefore, this method is workable, but not sufficiently substantiated. Indeed, the speed limit is required to be filled with a reinforcement constraint, which creates additional difficulties in solving this problem. The negative coefficient in the derivative link also causes some doubts, although in general the method works.

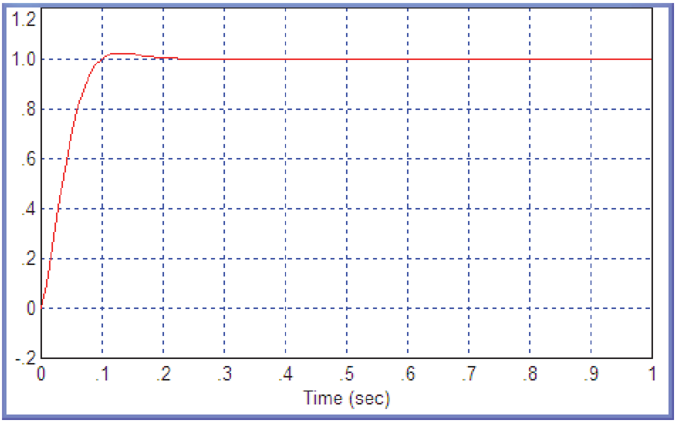


Fig. 6.20. The transient process in a system optimized according to the structure of *Fig*. 6.19 with an integration step of 0.01 *s*

6.6. Forced limitation of the frequency range of the model used in optimizing of the regulator

A developer can forcefully limit the frequency range obtained as a result of optimization of the system due to the fact that an additional inertial link, filter or delay link is introduced into the existing model of the object, which begins to affect only at those frequencies, at which he has not reliable data on the phase shift in the real object model. This method is based on the assumption of the worst situation. Optimization of the regulator for such a deliberately degraded version of the model of the object will give such a regulator, which with a high degree of probability will successfully work in a system with a real object. Fig. 6.21 shows a modified structure for this purpose to optimize the regulator. This method is patented and its effectiveness is proved by modeling.

Example 26. Let us solve the problem of Example 25, which consists in optimizing the PID regulator for the object (6.1). We use the structure of Fig. 6.21, which for the given problem takes the form shown in Fig. 6.22. We specify the delay in the element that limits the speed of the object, at the level of 0.05 *s*. Fig. 6.23 shows the obtained transient process, and the resulting coefficients, according to Fig. 6.22, are equal to $k_p = 1380$; $k_i = 11.2$; $k_d = 23.19$. The resulting transition process is shown in Fig. 6.23 (line 1).

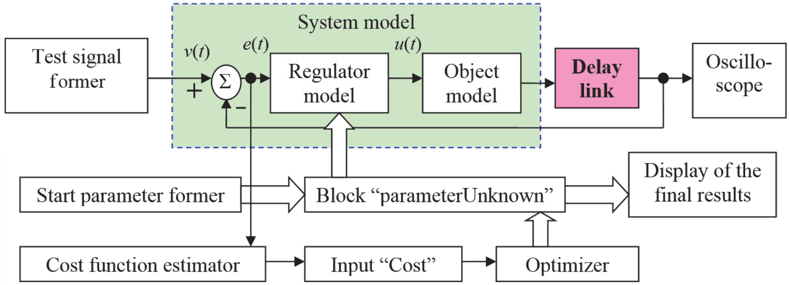


Fig. 6.21. Modified structure for optimizing the feedback loop regulator

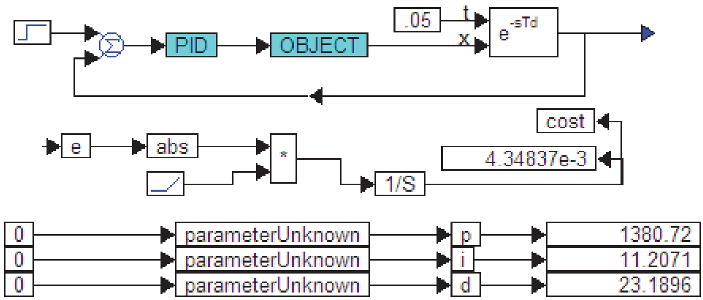


Fig. 6.22. A modified block diagram for optimizing the regulator in Example 25, with the object performance limited by the method of Fig. 6.21

For comparison, we can consider the process in the absence of delay (Fig. 6.23, line 2). It is also possible to consider a family of transient processes when the delay is varied from 0.2 *s* to 0.6 *s*, as Fig. 6.24 shows.

Based on this example, we can conclude that the artificial limitation of object model speed is effective for designing a regulator by the method of numerical optimization. This method allows solving even such problems that are formally incorrect for this method.

In addition, this method allows us to supplement the model, which is not well known, in such a way that the result of the simulation is in sufficient agreement with the result of the practical use of this regulator. Indeed, if the model of the object is known in a limited frequency band, the hypothesis that beyond the limits of this frequency band there is a decisive action of the delay link, therefore, when optimizing the regulator, a model supplemented with such a delay is used.

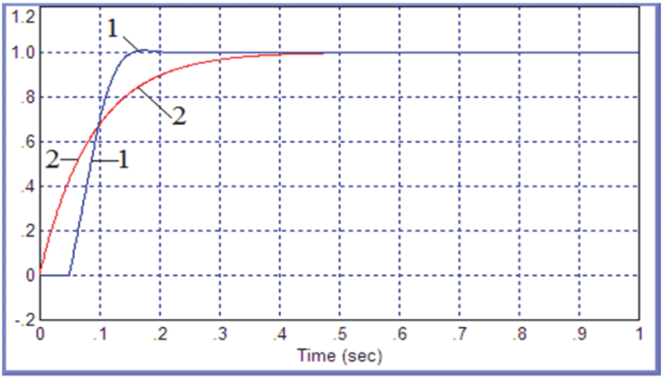


Fig. 6.23. The transient process in a system optimized according to the structure of Fig. 6.22 with an integration step of 0.01 *s* (line 1), as well as a process with no delay (line 2)

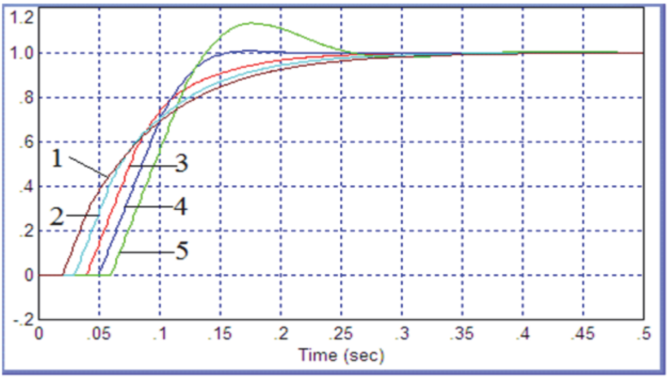


Fig. 6.24. Transient processes in a system optimized according to the structure of Fig. 6.22 with different delay values, the delay value can be determined from the beginning of the graphs: the line 1 is 0.2 *s*; line 3 is 0.4 *s*; line 3 is 0.5 *s* and so on

If the actual object has such a delay, the system will be stable, if the object does not have a delay, or it is less, the system will also be stable. Finally, if the frequency band in the object is limited not so much by delay, as by a higher-order filter (affecting this frequency band), then in this case the received regulator will provide stable control with a sufficiently high quality (i.e., with a slight overshoot).

7. DIVIDING OF THE MOTION: THE USE OF MULTIPLE DRIVES (MISO)

7.1. Justification for an excessive number of the influencing actions on the object

As a rule, the number of controlled quantities is the same as the number of signals that act. However, the use of an excessive number of inputs to an object is in many cases not only justified, but also necessary. If one drive (that is, an impact device on the output value of an object) allows to control an object at a high speed, but in a small range, and the other drive allows to control in a large range, but at a low speed, then combining these drives can supposedly combine these dignity and eliminate shortcomings. The lawful is the formulation of the problem: it is necessary by means of two such drives suppress two types of disturbances. Namely, small high-frequency noise and large low-frequency noise. Such a symbiosis of drives cannot cope only with large high-frequency noise, but such interference may not be due to the nature of the object. Thus, the two drives in the above example are not at all redundant, since their action is frequency-spaced. This example is one of the most important reasons for using the method of separation of movements.

7.2. Combining of the advantages of the different drives

With the help of motion division it is possible to provide control in a wide band and at the same time with a large range simultaneously by using a broadband drive with a small control resource and a narrowband drive with a large control resource.

Let consider the most general requirements for the regulator, as a consequence of the mathematical model of the object and the requirements for the open-loop transfer function. Fig. 7.1 gives the structural scheme of the system in general form. In the scheme: $W_{\mathrm{R}i}$ are the transfer functions of the individual links of the controller, $W_{\mathrm{O}i}$ are the transfer functions of the individual links, W_{OL} is the connection of all links of the object (and the transfer function of the open loop), $H(s)$ is the interference transformation $h(t)$ normalized to the object output, $N(s)$, $V(\mathrm{s})$, $E(s)$, $Z(s)$, $Y(s)$ are the transformations of noise $n(t)$, the prescription $v(t)$, the error $e(t)$, the signal from the sensor output $z(t)$ and the output signal $y(t)$.

From the structural diagram, the following relationships are obvious:

$$Y = \frac{W_{OL}}{1 + W_{OL} W_{FB}} \cdot R - \frac{W_{OL}}{1 + W_{OL} W_{FB}} \cdot N + \frac{1}{1 + W_{OL} W_{FB}} H \ . \qquad (7.1)$$

With unit feedback $W_{FB} = 1$, this expression is simplified. It is expedient to supplement the transfer function of the sensor with an inverse transfer function (as far as possible, as minimum, in the low-frequency region), so that the product of these two functions would be close to unity. This in practice is called "unit feedback". By introducing the notation of the corresponding partial transfer functions by errors, we obtain

$$Y = W_V \cdot R - W_N \cdot N + W_H \cdot H \ . \qquad (7.2)$$

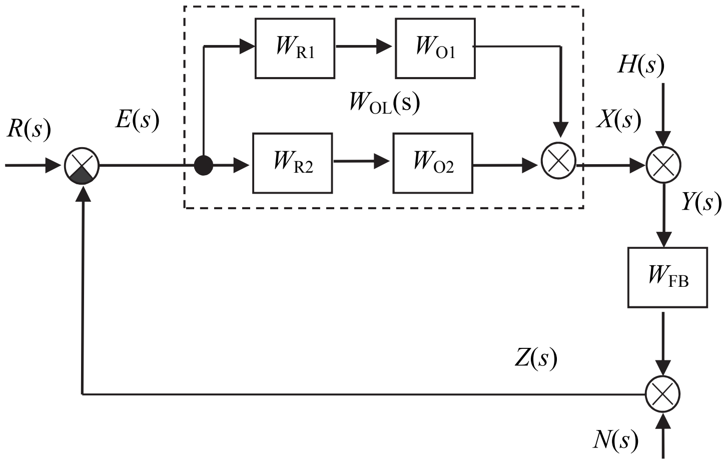


Fig. 7.1. Structural diagram of a system with two drives

It is seen that for $W_{FB} = 1$, an increase in the open-loop transfer function W_{OL} approximates the transfer functions indexed by R and N to unity, and this indexed by H to zero. The noise of the measurement generates a noise of the stabilized value. The disturbance can be suppressed to a value commensurate with noise, if the dynamic characteristics of the object allow providing the required band. It follows from this that the more W_{OL}, the better provided the stability of the entire loop. This is true for all frequencies within the bandwidth of the ω_C system. The consequence of this is the assertion that the more steeply the W_{OL} chart grows when moving from the higher frequencies to the lower ones, the better. However, to ensure stability in the area of unit gain, this slope should be kept unit (–20 *dB / dec*). Therefore, in the low-frequency region, developer can require a double slope, if it does not violate the stability of the system. Hence, the desired form of LAFC (i.e. AFC in the logarithmic axes) has a section of a unit slope near the unit gain (i.e., with $L(\omega_C) \approx 0$). In the region $\omega < 0.02\ \omega_C$, it is advisable to increase the slope up to the second order, and the length of this section can be several decades (see Fig. 7.2).

In fact, there are two control loops in the system: the first (fast) loop is formed by the transfer functions W_{R1}, W_{O1} and W_{FB}, the second (slow) loop is formed by the transfer functions W_{R2}, W_{O2} and W_{FB}.

The task to design the regulators W_{R1} and W_{R2} for this structure is as follows: it is required to provide the best LAFC with available resources in the form of transfer functions of fast and slow modulators W_{O1} and W_{O2} and with the existing limitations on their capabilities.

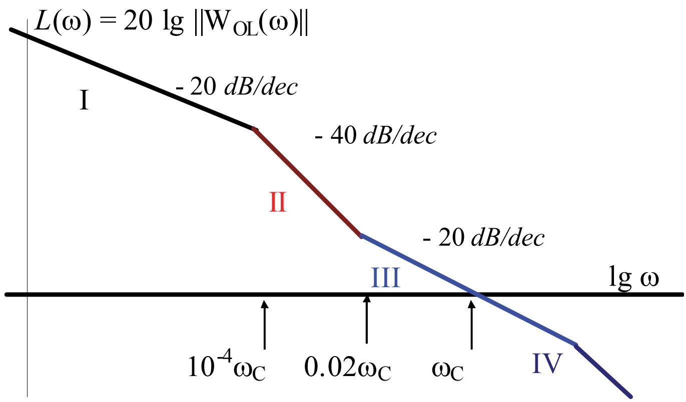


Fig. 7.2. The desired form of the logarithmic response (LAFC) of the open loop $W_{OL}(s)$

Based on known stability criteria for the closed loop of the automatic control system, one can formulate the requirements for logarithmic AFC. Slow and fast loops, are formed by the transfer functions $W_1 = W_{R1}W_{O1}$ and $W_2 = W_{R2}W_{O2}$.

The low-frequency section I (see Fig. 7.2) provides high accuracy in the low-frequency region because in this range the value of LAFC it is large and additionally increases with decreasing frequency. The mid-frequency section III ensures the stability of the system due to a unit slope in the region where the LAFC is close to unity. The high-frequency section IV has a large slope due to the fact that we cannot make a smaller slope in this region due to physical limitations on the capabilities of the object and (or) the sensor. The conjugating section II ensures, as soon as possible, the fastest possible attainment of large values of the LAFC as the frequency decreases, that is, when moving from section III to section I. The increasing of the steepness of this section allows increasing the accuracy and extending the range in which the error is less than the specified value. However, the excessively large steepness of this section threatens to lose stability of the system due to a large phase shift, which affects the phase shift of section III.

Two modulators were necessary in connection with the fact that it is not possible to provide LAFC with one modulator, which is shown in Fig. 7.2.

A fast modulator cannot provide a large control range, so it is not possible to realize low-frequency section I with it. The slow modulator cannot provide wide-band control, so it is impossible to realize high-frequency section III with it.

If we just try to turn on the two modulators in parallel, this is not going to help, since the control signals are distributed equally between the paths. For low-frequency signals the fast modulator enters the limit mode and ceases to participate in the control. This will lead to loss of stability of the system, or stability will have to be provided through the operation of a slow modulator, and therefore the band of operating frequencies will be determined by them what is insufficient.

To solve this problem, it is required that the first section would be formed by a slow modulator, and section III would be formed by a fast modulator. Thus, it is sufficient in section I to provide $W_1 >> W_2$, and in section III to provide $W_1 << W_2$. To do this, the specified graphs must intersect at a certain angle, and not be parallel. Graphs of LAFC slow and fast loops should be located, as shown in Fig. 7.3.

Regarding to the proposed procedure, the calculation of the specific values of the transfer functions of the regulator W_{R1} and W_{R2} (according to the known W_{O1} and W_{O2} without taking into account the dynamic nonlinearity) does not present any problems and can be performed graphically using LAFC. Analytical calculation taking into account the nonlinearity is difficult, however the numerical optimization of the regulator with known parameters of the model also does not present significant problems if taking into account the capabilities of software *VisSim* 5.0/6.0/7.5 and the available arsenal of methods and cost functions [21–42] for optimizing regulators.

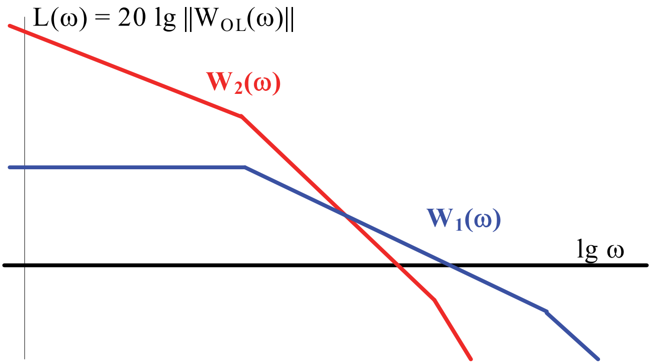


Fig. 7.3. Desired form of logarithmic fast and slow loops

The structure shown in Fig. 7.1 seems more complicated for analytical research than the traditional single-loop, but numerical optimization of regulators with the choice of this structure is complicated slightly.

Example 27. Let consider an object having two control channels. Each channel is described by its transfer function, and the output signal of the object is the sum of the outputs of each transfer function:

$$Y(s) = W_1(s)U_1(s) + W_2(s)U_2(s)\,, \tag{7.3}$$

$$W_1(s) = \frac{\exp(-s)}{100s^2 + 10s + 1}\,, \tag{7.4}$$

$$W_2(s) = \frac{1}{10s^2 + 1s + 1}\,. \tag{7.5}$$

In addition, we assume in the second channel a restriction in the interval from –1 to +1. One could suggest using only one channel to control the output signal. It is natural to choose the best channel. However, under these conditions, each channel has drawbacks and advantages, namely: the first channel is slower; the second channel has a smaller control range, which can affect the change in the prescribed value of the output value by an amount exceeding the level of limitation, for example, 10 units.

First, let us try to use only one control channel. Fig. 7.4 shows a block diagram for using only the first channel. The resulting transient process is shown in Fig. 7.5. This process is characterized by an excessively large overshoot at about 70%.

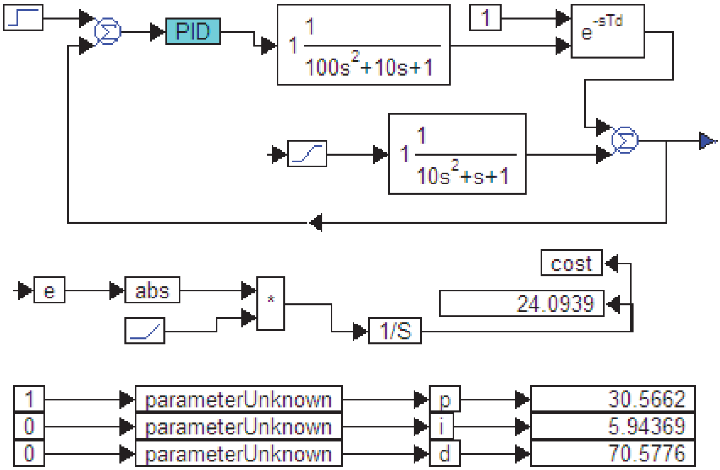


Fig. 7.4. The block diagram for optimizing of the regulator only on the first control channel in Example 27

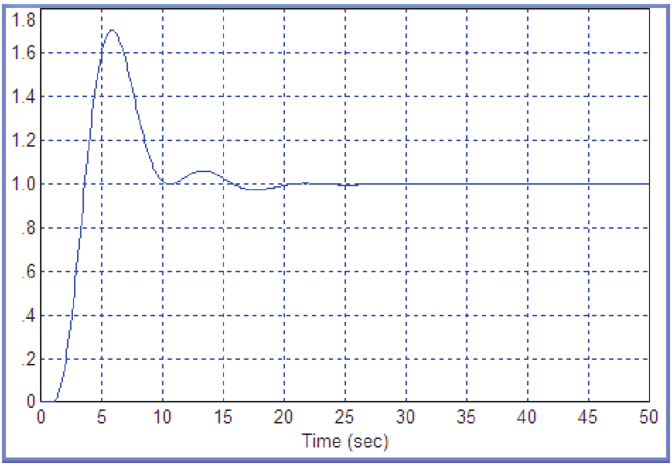


Fig. 7.5. The resulting transient process in a system with a regulator optimized according to the scheme in Fig. 7.4

Fig. 7.6 shows the scheme for optimization of the regulator only on the second channel. The resulting transient processes is shown in Fig. 7.7. This process is characterized by high quality but the second channel can implement control only in the range from –1 to +1.

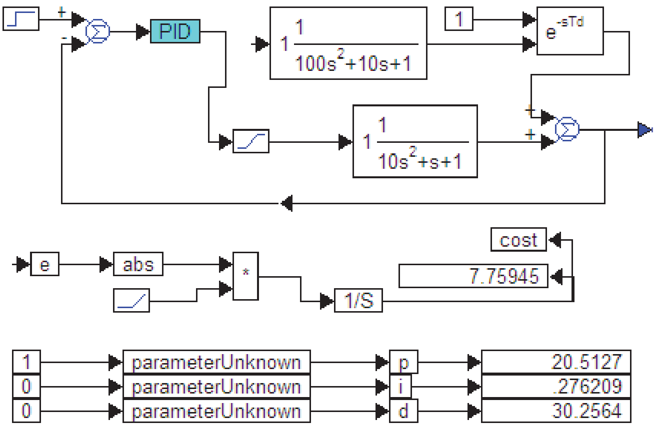


Fig. 7.6. The block diagram for optimization of the regulator only on the second control channel in Example 27

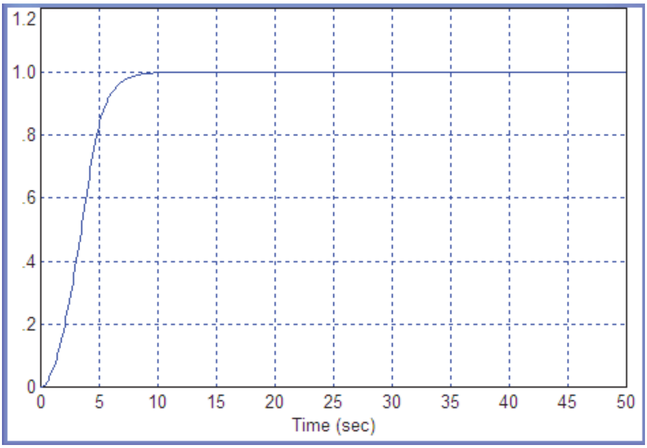


Fig. 7.7. The resulting transient process in a system with the regulator optimized according to the scheme in Fig. 7.5

This is not enough, so it is advisable to use the both control channels.

Fig. 7.8 shows the structure for optimization the two PID regulators, respectively, for the first and second control channels, and the resulting transient process when the jump is executed by a value of 10 units is shown in Fig. 7.9. The resulting transient process is characterized by the same high quality as the process of controlling the second channel, but the output signal value varies within a range that can only be achieved with the first channel. Thus, both channels contribute to the successful control of this object.

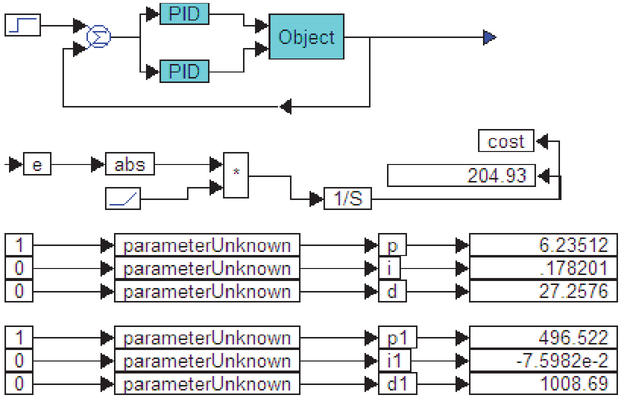


Fig. 7.8. The block diagram for optimizing the regulator only on the second control channel in Example 27

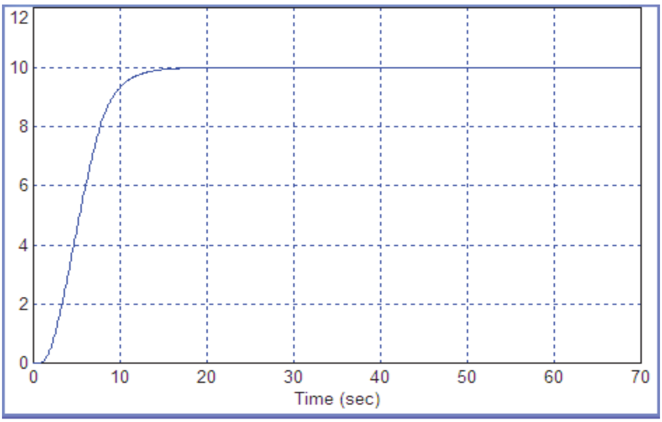


Fig. 7.9. The resulting transient process in a system with a regulator optimized according to the scheme in Fig. 7.8

Fig. 7.10 illustrates the output signals in each channel of the object individually. Fig. 7.11 shows the control signals in each channel separately.

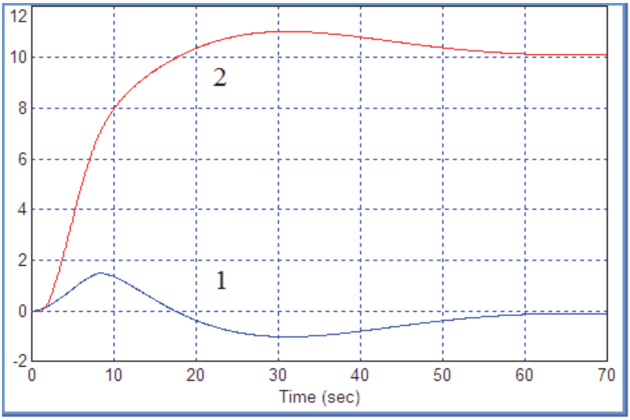


Fig. 7.10. Separate transient processes in each channel of the object, separately from the system with a regulator optimized according to the scheme in Fig. 7.8: line 2 is the first channel; line 1 is the second channel

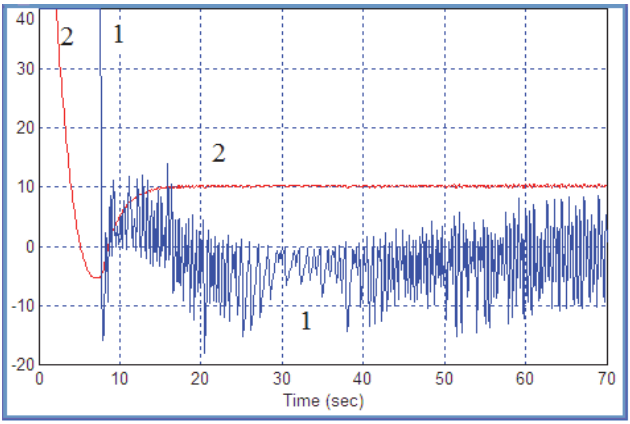


Fig. 7.11. Separate transient processes in each control channel of the object separately in a system with a regulator optimized according to the scheme in Fig. 7.8: line 2 is the first channel, line 1 is- the second channel

8. DIVIDING OF MOTIONS: USING OF MULTIPLE SENSORS (SIMO)

8.1. Combining of the advantages of the different sensors

Sometimes it is necessary to use parallel-connected sensors of the same controlled value, based on different principles or having different design or metrological characteristics. The seemingly excessive number of input sensors is in fact not excessive, because they have different frequency ranges of their effective action.

Often, measuring a signal in a wide band is characterized by a large average shift, and precision measurement methods require long averaging to reduce the effect of noise. Regulators based on the method of dividing of motion (MDM) allow combining of the advantages of low-noise unstable sensors, on the one hand, and stable noisy sensors, on the other.

In a number of cases, it is not possible to measure the stabilized value qualitatively in a wide frequency range: broadband sensors are characterized by large displacement or drift, and sensors with a small offset form noisier signal. Therefore, for example, things are in the transportable frequency standard [21, 22]. The use of MDM regulators allows combining advantages and eliminating the disadvantages of such sensors, i.e. creating a system with a small offset and low noise level, although none of the sensors does simultaneously produces signal with the specified quality. The idea of the method is illustrated in Fig. 8.1, although this structure is not recommended for use.

This figure shows the models of two sensors of the same value $Y(t)$, these sensors has different types and different advantages and disadvantages. If one of the sensors has only merit in relation to the other, without having the disadvantages, then the other sensor is not necessary. Therefore, it is advisable to consider such structures only if not all the advantages are concentrated in the properties of one sensor but they are distributed among the sensors. For example, suppose for certainty that the sensor 1 has a high accuracy with respect to a constant error value (and in the low frequency region), but its accuracy in the high frequency region is poor in comparison with the sensor 2. This can be manifested in two ways: in *variant A*, there is no signal at high frequency; in *variant B* it contains a large number of noises. Sensor 2 has complementary advantages and disadvantages. Namely, it provides high accuracy (or high signal-to-noise ratio) in the high-frequency region, but has low accuracy in the low-frequency region, for example, has a zero-level constant bias (*variant C*) or does not respond at all to a constant level (*variant D*).

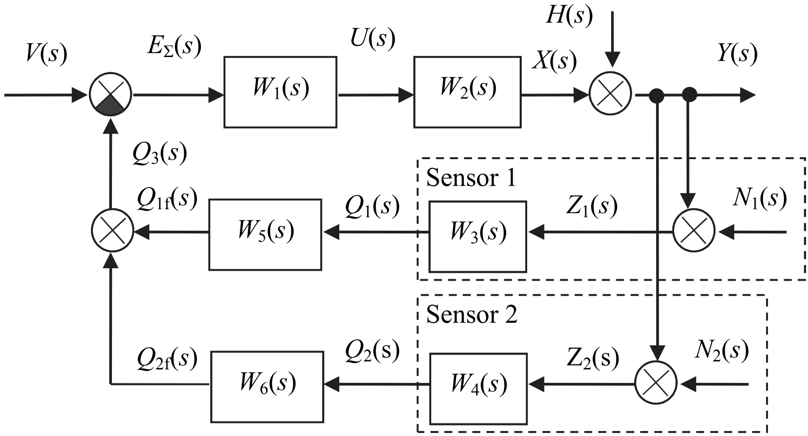


Fig. 8.1. Not recommended block diagram of system with two sensors

All these options can theoretically be reduced to the "*variant A + D*", that is, to ensure that the system works in such a way that the sensors produce a signal only in the area of their best work, and in the area of their worst operation their signal is zero. To do this, it is sufficient to calculate and apply the corresponding filters at the outputs of each of the sensors, W_5 and W_6 in Fig. 8.1.

If it is possible to ensure such matching of these filters, that their boundary frequencies coincide, and the order will be the same, it can be hoped that there will be three frequency bands with the following properties:

a) in the low-frequency range, the transfer function W_3W_5 is close to unity, the transfer function W_4W_6 is close to zero;

b) in the low-frequency range, the transfer function W_3W_5 attenuates from unity to zero; the transfer function W_4W_6 increases with the same speed from zero to one, in the sum these functions in the whole range make up unity;

c) in the low-frequency range, the transfer function W_3W_5 is close to zero, the transfer function W_4W_6 is close to unity.

In this case, the block diagram of Fig. 8.1 can be applied relatively successfully. If this requirement is not ensured, then such a structural scheme is not completely correct. This structure is operable only if the combined value $q(t) = q_{1f}(t) + q_{2f}(t)$, composed of the sum of the output filtered signals of the two sensors, is close to the output controlled value $y(t)$.

Therefore, the structure in Fig. 8.1 is not recommended for use, although it clearly demonstrates the principle of the system with two sensors. To make it clear that the correctness of this structure depends on the properties of the filters, let us imagine that the transfer functions W_3W_5 and W_4W_6 are equal to one in the whole range, and the noise of both sensors is equal to zero for simplicity, that is, we consider the case of two ideal sensors in this scheme. Then the signal applied to the subtractor is equal to twice the output signal of the system, that is, there is a coefficient equal to two in the feedback. In this case, if the system is stable, the output signal is equal to one-half of the prescription: $y(t) = 0.5\ v(t)$. Thus, the structure shown in Fig. 8.1, requires accurate matching of the filters and transfer functions of the sensors in the entire operating frequency band, where such matching is violated, the output signal repeats the task not exactly, but with the inverse of the total feedback transfer function.

Let consider the structure in Fig. 8.2. In this structure, the errors of $y(t)$ are calculated on the basis of the output signals of two sensors: $q_1(t)$ and $q_2(t)$. Each of the errors received must be zero. If the feedback signal equal to the sum of these errors is fed to the object with gain through the regulator W_1, then the total error is reduced to zero. For the correct operation of this circuit, it is not required that in the section "b" the transfer functions W_3 and W_4 smoothly pass into each other. It is enough that the failure between them is not significant, it is preferable that the sum of these functions exceed unit value in comparison with the failure variant. If both sensors are ideal even in the absence of filters, that is, with $W_3 = W_4 = 1$, $N_1 = N_2 = 0$, we obtain $e_\Sigma(t) = 2e(t)$. This is equivalent to doubling the regulator factor. If the system remains stable, it works correctly, and the output signal in the working frequency region repeats with the required accuracy the reference signal $y(t) \approx v(t)$.

However, in the considered structure according Fig. 8.2, filtering of sensor signals is not applied (transfer functions W_3 and W_4 refer to sensor models). Therefore, it is advisable to supplement each sensor path with an appropriate filter. The requirements for these filters are greatly simplified in comparison with the requirements for filters in the structure of Fig. 8.1.

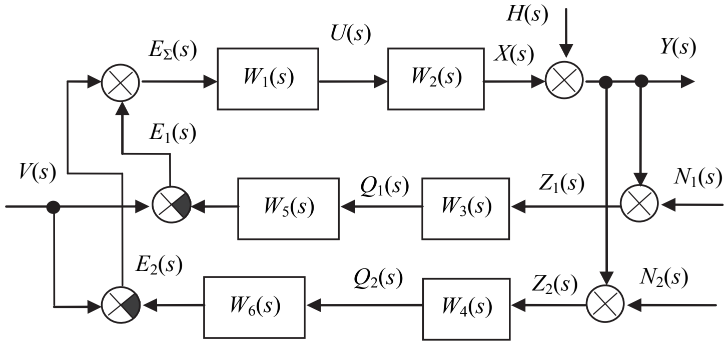


Fig. 8.2. Recommended block diagram of system with two sensors

They are as follows:

1. In the frequency range corresponding to the best accuracy of the *i*-th sensor, the transfer function of the $W_{i+2}W_{i+4}$ path should be significantly (5 or more times) greater than the transfer function of the other path containing the sensor with the worst accuracy.

2. In the transition region, the transfer functions of both paths must be mated so that their sum is always greater than one.

If these requirements are met, the total error signal will always be equal to or greater than the true error, which does not contradict the principles of the system operation.

For the convenience of design, the transfer functions of the filters can be transferred through subtracting devices, each of these functions is multiplied by the transfer function of the regulator, eliminating a separate block with this transfer function. In this way, we obtain a system with two sensors and two regulators, shown in Fig. 8.3.

Example 28. Consider the object from Example 23. Suppose that the output value of an object is not available for measurements directly. Let the first sensor have a constant offset at 0.2 units, and the second sensor has a Gaussian noise with a variance of 0.2 units. We organize a two-channel PID regulator for controlling this object on two sensors simultaneously. The corresponding structure is shown in Fig. 8.4. Let us make some specifications. Since the first sensor has an offset, the signal from it is not suitable for integration, so instead of the PID regulator in the first channel, we must use the PD regulator. For this purpose, we set the coefficient of the integrating link in the first regulator to zero. Since the second sensor is characterized by strong noises in the high frequency region, it is inappropriate to use derivation in the second regulator, so we will use only the PI regulator for the second sensor. For this goal, we set the coefficient of the derivative link of the second regulator to zero. The transient process in the resulting system is shown in Fig. 8.5. The steady-state value in the system is zero; the effect of the noise of the second sensor is not significant. Fig. 8.6 shows the results of measurements of the output signal by two different sensors. For comparison, we will simulate the system only with the first sensor or only with the second sensor. The results are shown in Fig. 8.7 and Fig. 8.8. In the same box, the received coefficients of the regulators are shown.

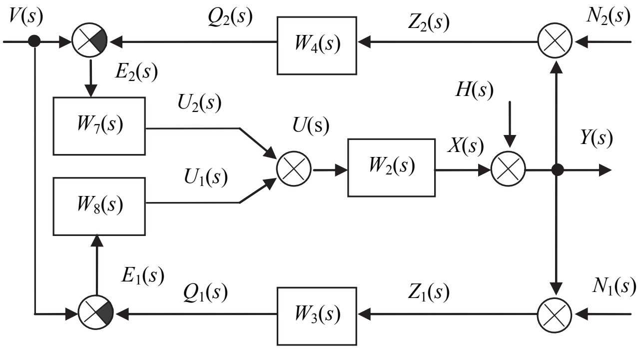


Fig. 8.3. Recommended block diagram of system with two sensors

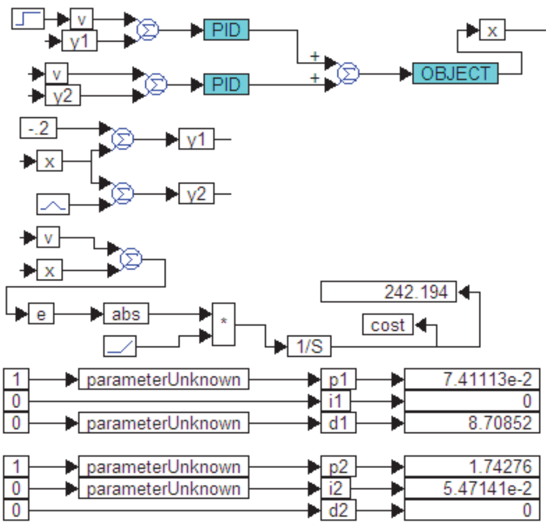


Fig. 8.4. A block diagram for controlling a system with two sensors from Example 28

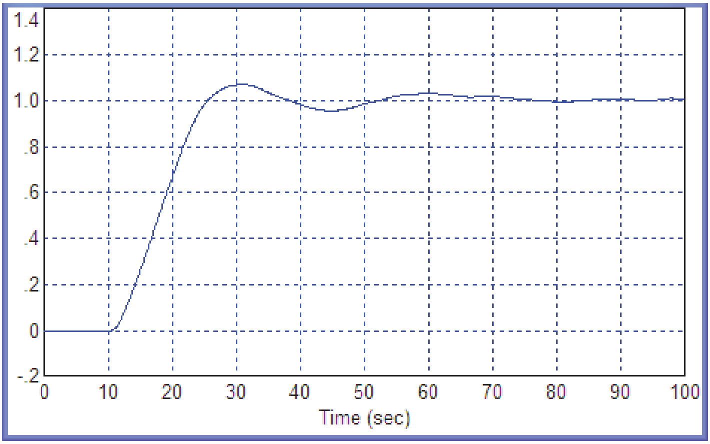


Fig. 8.4. A block diagram for controlling a system with two sensors from Example 28

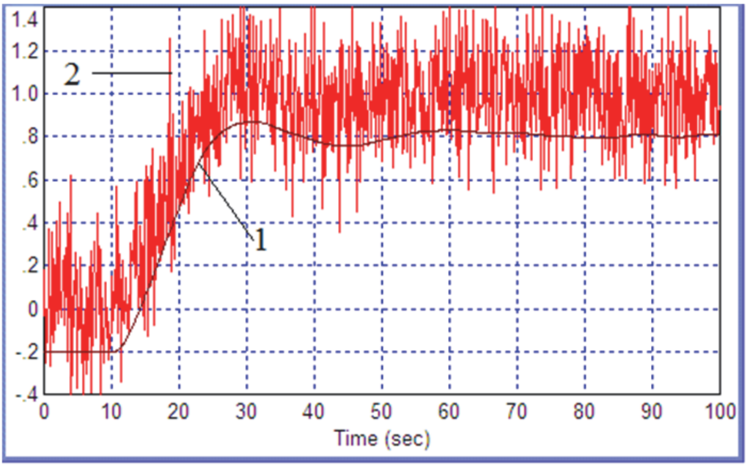


Fig. 8.6. The transient processes simulating the sensor output signals in the system of Example 28 in the structure of Fig. 8.4 (line 1 is sensor with displacement, line 3 is sensor with noises)

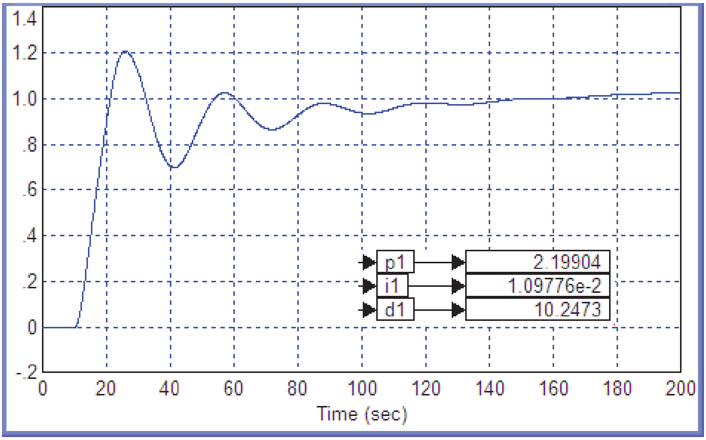


Fig. 8.7. The transient processes in the system from Example 28 in the structure of Fig. 8.4

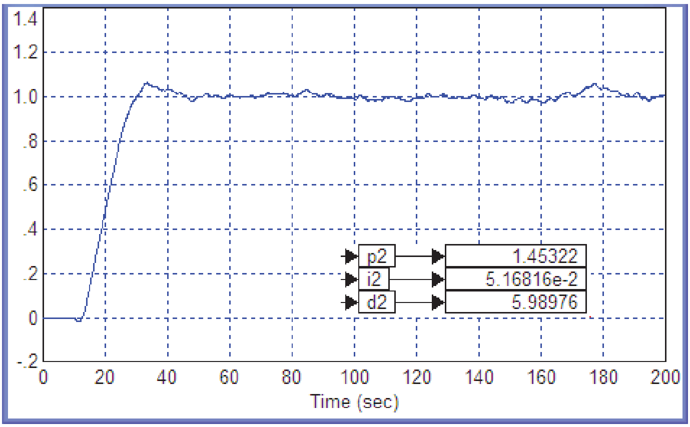


Fig. 8.8. The transient processes in the system from Example 28 in the structure of Fig. 8.4

Comparing the results with the results obtained using only one channel. It is seen that the offset remains with the first sensor in the system. The resulting integral path coefficient in the first case is extremely small, in the system the transient process lasts longer than the simulation time is set, which makes it possible to reduce the error at the end of the simulation interval. Thus, it is confirmed that when using a displacement sensor, the system has an appropriate static error. In a system using only the second sensor, the output noise is greater than in the system using both sensors.

8.2. Simultaneous combing of the advantages of the different sensors and different drives

Both of the principles discussed above, the separation of movements by sensors and the separation of motions by modulators (drives), can be used simultaneously if these problems occur both with respect to sensors and with respect to modulators. This will create a system that measures one output signal using two or more independent sensors and controls the object using two or more modulators. However, in spite of the fact that the regulator will have two inputs and two outputs at the same time, **such a system will not be multichannel**, since the control will be performed only with respect to a single output value.

9. MODIFICATION OF THE COST FUNCTIONS

9.1. Providing of the power saving

As discussed in Section 5, in order to optimize a feedback system, the structure shown in Fig. 5.6 can be used. This structure can be generalized and simplified by the structure shown in Fig. 9.1. Here the cost function calculator can be generally called the “signal analyzer”.

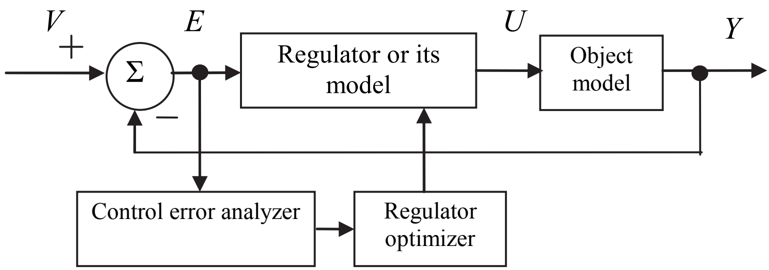


Fig. 9.1. Generalized and simplified structure for system optimization

In this case, signal analyzer is connected only to the output of the subtractor, which calculates the control error. If a composite cost function is used, additional inputs of this analyzer can be used, connected to other points in the system model.

For example, to provide energy savings in the cost function, an integral of the square of the control action can be introduced, which represents the generalized control energy. The corresponding structure is shown in Fig. 9.2. In fact, the role of the signal analyzer is played by two blocks: a control error analyzer and an energy cost analyzer. The signals from the outputs of these blocks are fed to the optimizer, where they are added together with the weight coefficients. The adder can be selected from the optimizer structure and shown in an explicit form, which will formally form a new structure, but the essence of it will remain the same.

This structure of the model for optimizing the feedback system works as follows. The simulation software *VisSim* performs a multiple simulation of the action of the specified structure. At the input of the structure, which is the positive input of the subtractor, the input signal *V* is fed in the form, for example, of stepwise action. The negative input of this subtractor receives the output signal of the model of the object *Y*, which is also the output signal of the structure. The subtractor calculates the difference of these signals, called error E. The error is converted by the regulator into the control signal U, which is input to the object model input. If the error is positive, the control signal affects the model of the object so that the output signal of the object model increases, and the error thereby decreases. If the error is negative, then the control signal affects the model of the object so that the output of the object model decreases, and the error for the account of this increased.

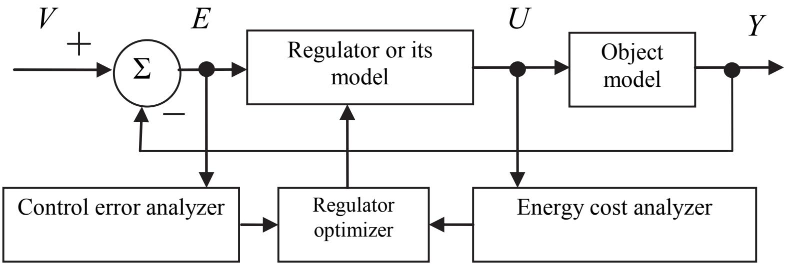


Fig. 9.2. Generalized and simplified structure for system optimization

If the error is zero, the control signal affects the model of the object so that the output signal of the object model does not change, and the error remains zero. The regulator optimizer generates the starting values of the regulator coefficients according to the given algorithm, analyzes the outputs of the analyzer of achievement of the control goal and the energy cost analyzer, calculates the cost function and forms on this basis new values of the regulator coefficients. Using any of the methods of multidimensional optimization (for example, the *Powell* method) the optimizer searches for such values of the regulator coefficients that provide the minimum value of the cost function. These coefficients are the result of applying the structure of the model to optimize the feedback system. They are extracted from the optimizer's memory and used in the production of the feedback system. If the model of the object is determined quite accurately, then the system works with the same quality and accuracy indicators that were achieved in modeling using the described structure of the model. Thus, the task is solved. If the model of the object is determined only in a limited frequency range, an element with a limited speed, for example, a delay link, as suggested in Section 6, can be used, see Fig. 6.21.

Energy-saving regulators are particularly effective in saving energy costs for control if the integrator is included in the object model. Indeed, in this case, energy is consumed only in the process of transferring an object from one state to another, and the stay of an object in some equilibrium state does not require the expenditure of control energy. An example of such systems is the transition of a satellite from one stationary orbit to another, and so on.

Example 29. Let us consider an object whose transfer function contains a second-order object, a delay link and an integrator:

$$W(s) = \frac{\exp(-4st)}{(10s^2 + 10s + 1)s}. \tag{9.1}$$

We will optimize the PI regulator, since the integrator is contained in the object, and it is not required in the regulator. The structural diagram of the object in modeling in *VisSim* is shown in Fig. 9.3. The structural diagram of the entire system is shown in Fig. 9.4.

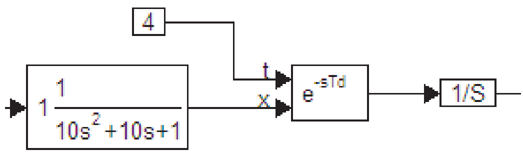


Fig. 9.3. The object structure for the simulation of Example 29

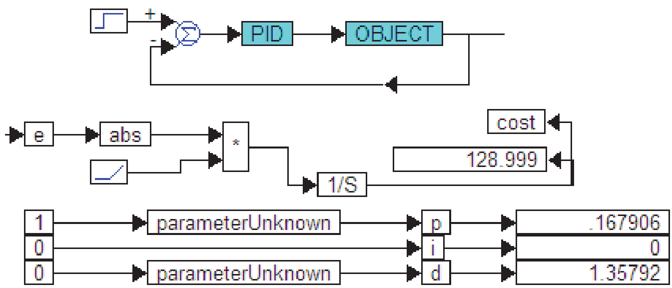


Fig. 9.4. The structure of the entire system from Example 29

Fig. 9.5 shows the obtained type of transient process and Fig. 9.6 is a graph of the control resource flow, which is the square of the control signal $U(t)$.

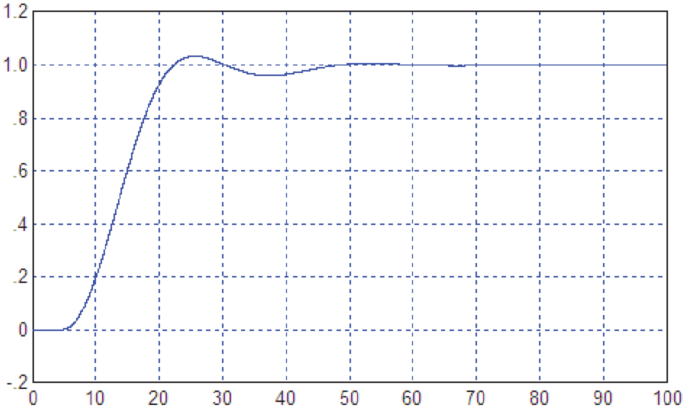


Fig. 9.5. Transient process in the system of Example 29

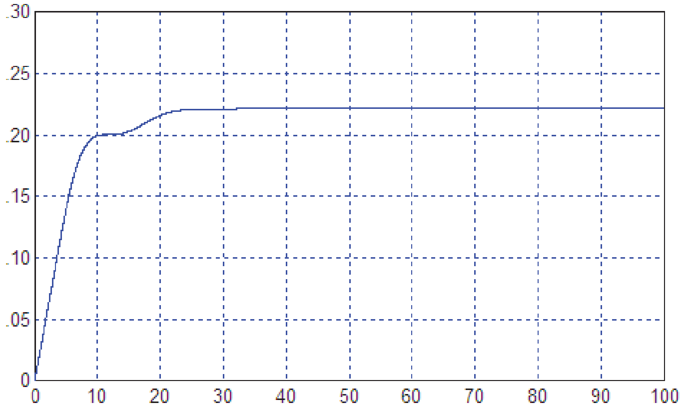


Fig. 9.6. The graph of the change in the flow of control resource according to Example 29

The total resource consumption for control is approximately 0.225 units. To design an energy-saving regulator, we introduce into the cost function the integral of the square of the control signal with a weighting factor. Simulation has shown that the weighting factor must be at least 100, so that energy saving works efficiently. We used a weight coefficient of 500. The structural diagram is shown in Fig. 9.7. The resulting transient process is shown in Fig. 9.8, and the graph of the change in resource consumption is shown in Fig. 9.9. It can be seen that the transient process not only did not deteriorate, but even improved, namely: the overshoot and the duration of the process were reduced from 50 to 30 *s*. The resource consumption became approximately 0.14 units, that is, 1.6 times less.

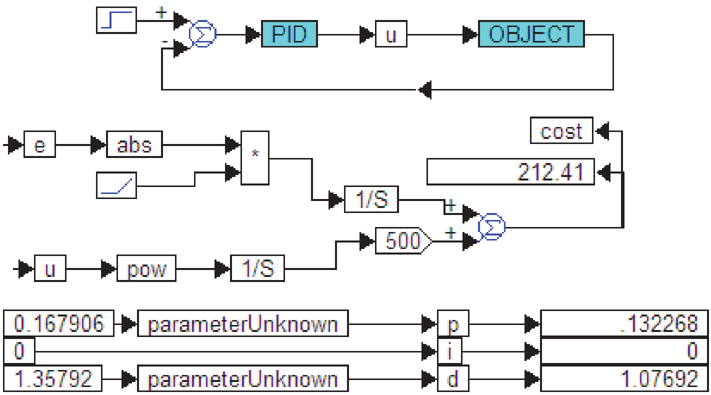


Fig. 9.7. Structure for optimizing of the energy saving regulator in Example 29

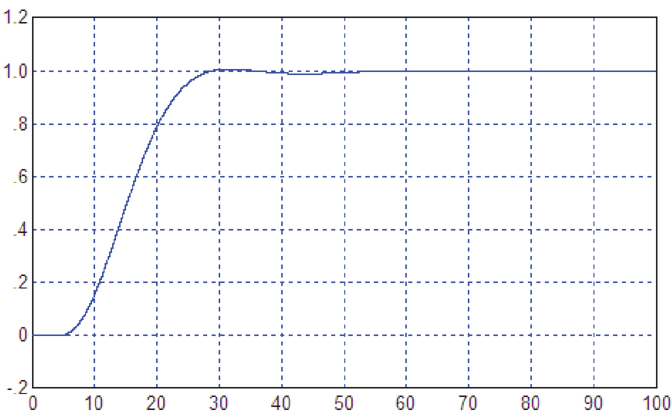


Fig. 9.8. Transient process in the energy-saving system of Example 29

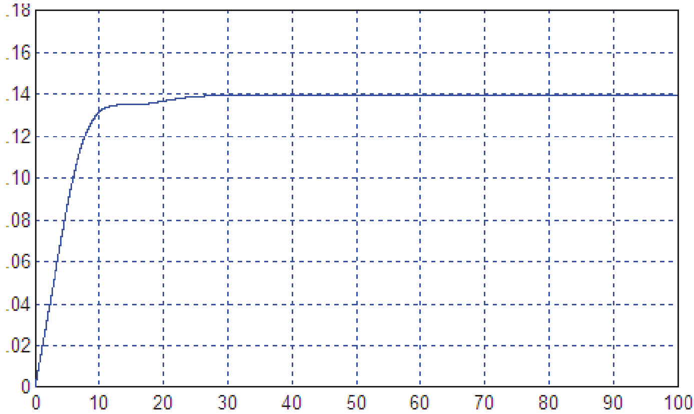


Fig. 9.9. The graph of the change in the flow of control resource according to Example 29

9.2. The detector of the error growth

In the objective function, we can also add the integral of the positive part of the product of the error to its derivative as a term. A developer should also use a weighting factor.

$$\Psi_2(T) = \int_0^T \{R[e(t)de(t)/dt] + |e(t)|t\}dt\,. \tag{9.2}$$

$$R[f] = \begin{cases} f, f > 0; \\ 0, f < 0. \end{cases} \tag{9.3}$$

Here R is the lower limiter that passes only the positive signal, and if the input signal is negative, but zero is output at its output.

The logic of this modification is this: if the error and its derivative coincide in sign, this means the wrong development of the transient process, the error increases when it should decrease or decrease when it should increase. The presence of such sections of the transient process is undesirable, such areas arise when the process has overshooting and (or) oscillations.

Example 30. Let us consider an object whose transfer function contains a second-order link and a delay link:

$$W(s) = \frac{\exp(-2st)}{10s^2 + 0{,}1s + 1}. \tag{9.4}$$

The model of this object is shown in Fig. 9.10. The structure of the system for optimization with the result obtained is shown in Fig. 9.11. The resulting transient processes are shown in Fig. 9.12.

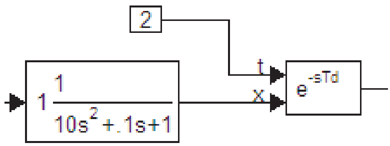


Fig. 9.10. Structure of the object for modeling from Example 30

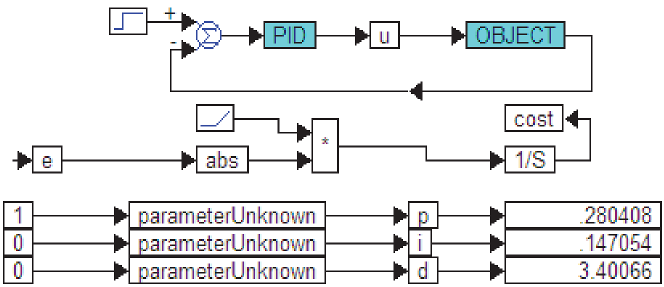


Fig. 9.11. The structure of the entire system from Example 30

We introduce the error growth detector in the cost function. In this case, only a weight function of at least 100,000 functions efficiently, as shown in Fig. 9.13. The resulting transient processes are shown in Fig. 9.14.

Fig. 9.15 shows the output signal of the error detector (before integration), which took place in the structure of *Fig*. 9.11. In the structure of Fig. 9.13 this signal is practically zero (in the same axes the remnants of this signal merge with the axis).

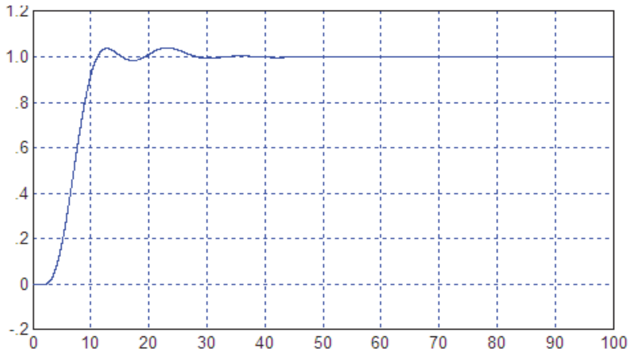


Fig. 9.12. Transient process in the system of Example 30

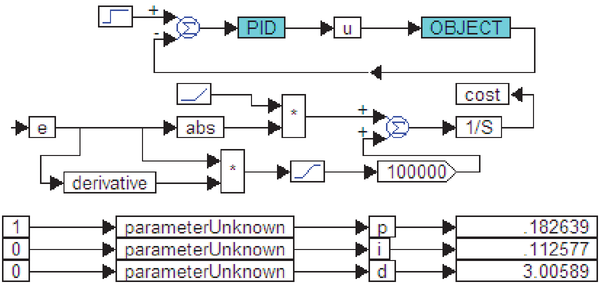


Fig. 9.13. The structure of the entire system of Example 30 using the error growth detector

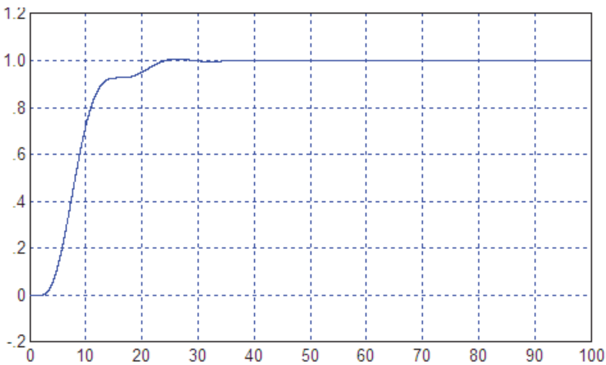


Fig. 9.14. The transient process in the system according to the structure of *Fig.* 9.13

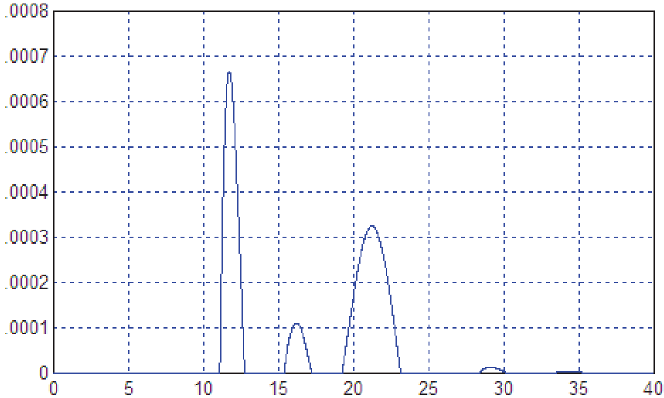


Fig. 9.15. The output signal at the output of the error detector delimiter (up to the integrator) in the system according to the structure of Fig. 9.11

9.3. Optimization with the use of etalon transient response

In some cases, even the use of all the methods discussed above does not allow us to calculate such a regulator that provides a transient process of the required quality. In particular, one of the most acute problems can be the reverse overshoot that occurs in the system.

The integral function of the square of the difference between the actual error $e(t)$ and the error of the reference model $e_R(t)$ describing the ideal process can be led to the objective function:

$$\Theta(T) = \int_0^T [e(t) - e_R(t)]^2 \, dt \, . \quad (9.5)$$

Fig. 9.16 shows the block diagram for optimizing the regulator by this method.

Here, the former of the reference process consists of a serial connected model of the object delay and the model of the ideal system. From the received ideal response of the system to the step jump of the prescription, the actual response is subtracted from the response with given parameters of the regulator, the difference is squared and goes to the input of the optimizer of the regulator. In this case, a cost function is used that contains, along with the other members, the integral of the square of the deviation of the actual transient process from its ideal model. The square of this deviation is minimized together with other positive components of the cost function. Such a method and such a structure make it possible to improve the quality of the transient process when other methods and structures do not give the desired result.

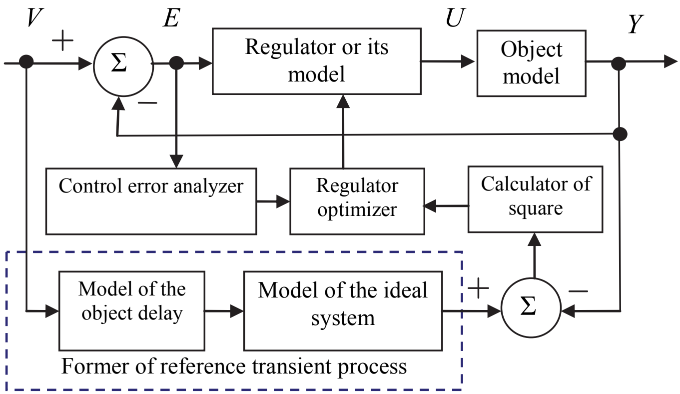


Fig. 9.16. Simplified structure for optimizing the system using a reference transient process

Comparison of the transition process with its ideal model is carried out only in the structure for numerical optimization, and not in the real control system. This fundamentally distinguishes this method, for example, from the method of controlling an object with a reference model. In the control method, the reference model is used during the functioning of the real system. Disadvantages of such a system with a reference model are not evident, but essential. Briefly, their essence lies in the fact that in real systems, testing the input signal is not usually the main problem, the most important task is to suppress the disturbance, and it is not available for measurement, so it is impossible to form an ideal response of the system. In modeling, this problem does not arise. In the adaptive system with the reference model, the difference in the output signal from its ideal model is considered for the operative modification of the regulator model. In the proposed structure, this difference is used to correct the regulator coefficients when it is numerically optimized, after which these coefficients are transferred unchanged to the regulator model of the real system and used without automatic tuning.

Example 31. Let us consider an object model in the following form:

$$W_4(s) = \frac{\prod_{i=1}^{4}(\tau_i s + 1)}{s\prod_{j=1}^{4}(T_j s + 1)} \cdot e^{-s}. \qquad (9.6)$$

This kind of transfer function can arise, for example, by approximating the logarithmic amplitude-frequency characteristic of a not integer slope with the help of several characteristics of an integer slope. In particular, such problems with such transfer functions arise as a result of the identification of objects with distributed parameters. Let us specify specific values for the coefficients in the numerator and denominator, namely: $\tau_1 = 100$, $\tau_2 = 10$, $\tau_3 = 1$, $\tau_4 = 0.1$, $T_1 = 300$, $T_2 = 30$, $T_3 = 3$, $T_4 = 0.3$.

We will develop a structure for optimizing the regulator for the object (9.5). The corresponding structure is shown in Fig. 9.17. This structure is developed in software *VisSim* 6.0. The structure contains compound blocks with the names “PI-Regulator”, “Object”, “Optimizer” and “Cost estimator”, respectively. Fig. 9.18 shows the model of the object in accordance with (9.5).

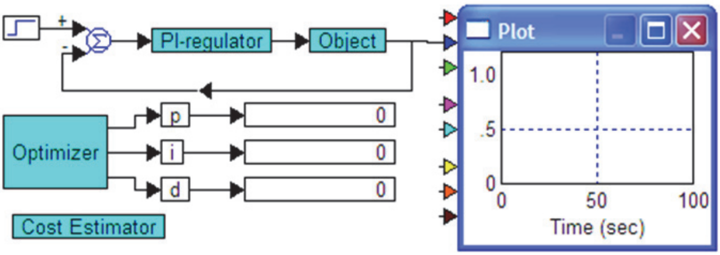


Fig. 9.17. System model in software *VisSim*

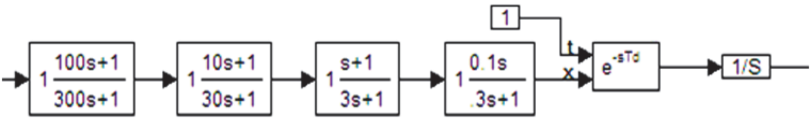


Fig. 9.18. Object model in software *VisSim*

Fig. 9.19 shows the PID regulator model. The block structure for optimization should contain the following elements: “parameter unknown” element, by the number of parameters to be found as a result of the optimization procedure, as well as constants specifying the initial search conditions. The view of this block is shown in Fig. 9.20. The values of the starting values are clear from the figure, these variables are, respectively, $K_p = 1$, $K_p = 0$ and $K_d = 0$.

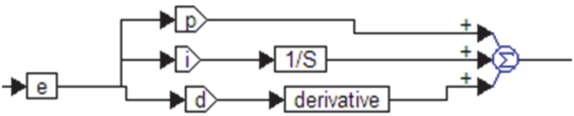


Fig. 9.19. The controller model in software *VisSim*

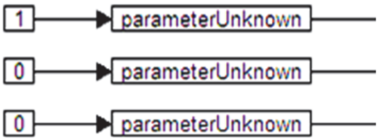


Fig. 9.20. Block model “Optimizer”
in software *VisSim*

Also, the “Cost” block must be entered in the model. It analyzes the cost function for calculating new values of optimized parameters. Also, in the structure, a cost function calculator is necessary, the structure of which is shown in Fig. 9.21. The output of this calculator must be connected to the input of the “Cost” element.

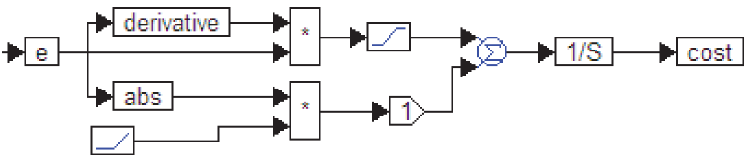


Fig. 9.21. Model of the “Cost Estimator” block in software *VisSim*

The cost function, which is calculated by the block shown in Fig. 9.21, contains two terms. The first term is the result of the product of the error on its derivative, this result passes through the limiter that cuts the negative part of this product, and leaves only its positive part. The output signal of this limiter is fed to the integrator through the adder, where it is integrated.

The second term is the product of the error by a linearly increasing function that simulates the time from the simulation start. At the output of the multiplier there is a scale transformation block introducing a weighting factor, which in this case is equal to unit. The result also passes through the adder to the integrator, where it is integrated. In this way, the block calculates the cost function with the required components.

Using the structure shown in Fig. 9.17 with the internal block structure shown in Fig. 9.18–9.21 gives the system, the transient processes in which are shown in Fig. 9.22. The line 1 shows the processes in the system, if in the value function only use the value function on the basis of the product of the error module for a time. The line 2 shows the processes obtained in the system, if we also use the term generated by the error growth detector to calculate the regulator.

In the first case, the resulting transient process in the system is characterized by a large number of oscillations. In this case, the first overshoot reaches 30%. In the second case, a process is obtained in which the overshoot is about 90%, the number of oscillations decreases sharply, as shown by the line 1 in Fig. 9.22. Therefore, in this case, adding to the cost function a term calculated by the error growth detector does not improve the result. The process shown by the line 2 is more attractive from the position of small overshoot, but less attractive from the perspective of the duration of the process. The introduction of the term (9.5) with the weight gives a system with a transient process, which is shown in Fig. 9.23.

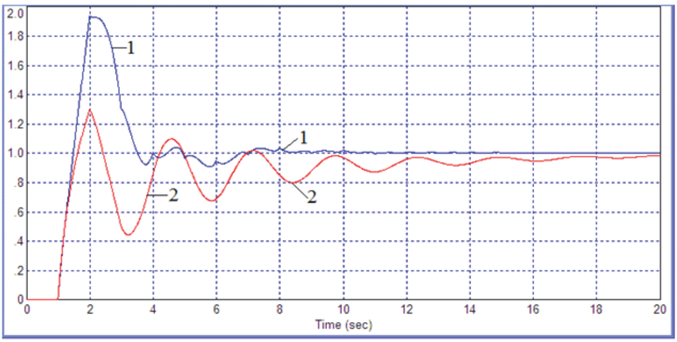


Fig. 9.22. Transient processes in a system with a regulator obtained for different cost functions: the line 2 shows the processes when using the objective function without the term (9.5) and without the error growth detector; the line 1 is the processes when using the cost function without the term (9.5) but with the error growth detector

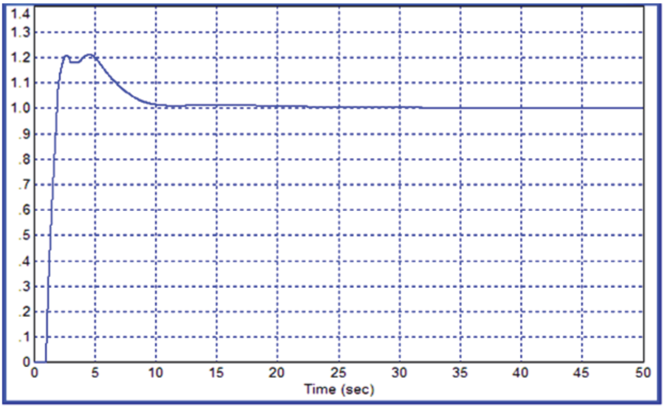


Fig. 9.23. The transient process in the system, obtained as a result of optimization with the cost function with the addition of the term (9.5) with the weight

Thus, the introduction of a square from the error module into the objective function made it possible to reduce the overshoot to 20%. The process is significantly improved. However, a further reduction in overshoot is not achieved with this structure when selecting cost function data, even if we change the weighting coefficients in some limit.

10. NEW STRUCTURES FOR SINGLE CHANNEL OBJECTS (SISO)

10.1. Robust power saving two-channel regulator with single output

Controlling of an object with one output value using two input effects is used quite widely, since it allows combining the merits of the two channels and overcoming their shortcomings, as shown above. This method can be used if there are special properties of the control object that are that the change in the output quantity can occur due to the combined actions of two factors, that is, the object can be controlled by two impact channels (two drives). The expediency of simultaneous use of the both channels is dictated by the fact that each channel has disadvantages, while the shortcomings of one of the channels do not coincide with the shortcomings of the other channel. Otherwise, a developer should use the best channel, and the worst one should not be used.

We proposed and investigated methods for calculating the regulator. In particular, it is shown that a developer can use software *VisSim*, and as a cost function it is advisable to use the integral of the error module multiplied by the time from the moment of the beginning of the transient process. It is also advisable to introduce additional terms with weight coefficients under the integral of the cost function, such as the square of the control action and others.

Of considerable interest is the use as an additional channel of control action, which is characterized by extremely limited capabilities, but there is a powerful argument for its use as a resource saving.

It is important to investigate how much such use can be justified, whether it is possible to provide robust control, which means in this case a fairly small effect of the accuracy of the calculated coefficients (as well as the accuracy of determining the parameters of the mathematical model of the object) on the stability of the system, and also on the type of the transient process.

These questions can be investigated by modeling, however, it is necessary to carry out a sufficient number of similar experiments with different models of objects, since some problems may not manifest themselves in some particular cases, but in other cases become insurmountable.

In this section we solve the problem of numerical optimization of the PD regulator for an object containing two control channels, each of which contains an integral component. In this case, each channel contains a second-order filter, and in the “worst” channel there is additionally a nonlinear element that converts a continuous signal into a discrete signal, at the output of which there is a derivative link. This makes the channel extremely inefficient in its pure form, which makes it urgent to investigate the feasibility of using such a channel in conjunction with the “best” channel, free of this nonlinearity.

Example 32. Let us consider an object whose mathematical model is given by the sum of two channels of influence. Let’s examine the numerical optimization method for calculating the regulator, based on a cost function that contains energy costs. In addition, we use a different cost of control resource with the same static transmission coefficient of each control channel. It is natural to assume that the price of the resource, which is used for managing the “worst” channel, is much lower than the price of the control resource on the “best” channel (the first one). Otherwise, the “worst” channel (second) cannot be useful. Since a criterion is used that includes resource saving, it is advisable to consider the class of objects containing the integrator. In the structure of the model, the integrator common to both control channels can be made for the adder.

Let the transfer function of the first channel has the following form:

$$W_1(s) = \frac{1}{0,1s^2 + 0,5s + 1} \cdot \frac{1}{s}. \tag{10.1}$$

Here s is the argument of the Laplace transform.

The transfer function of the second channel is given by:

$$W_2(s) = \frac{1}{s^2 + 2s + 1} \cdot \frac{1}{s}. \tag{10.2}$$

The transfer function (10.2) does not fully describe the second control channel. In order to describe its weaker control capabilities at the input of this channel, we introduce into the model of the second channel a nonlinear element and a derivative link with it in turn. This derivative link naturally compensates for the effect of the integrator; therefore, the effect on the second channel does not affect the rate of change of the output quantity, but only introduces almost stepwise changes in it, just as a discarded step at the time of its rejection can change the speed of the spacecraft due to the reactive effect. We specify the discreteness of the nonlinear element by a step size of 0.2 units. We set the ratio of management resource values as 1:50.

The method of solving this problem uses the numerical optimization by software *VisSim*. In this case, the structure of the regulator contains two channels, each of which consists of a proportional and derivative link, the outputs of which are summed. In the second path, we recommend that a developer would enter the "dead band" element, based on the results described above. The width of the "dead band" is commensurable with the discreteness of the nonlinear element of the second channel in the model of the object, but it is impossible to achieve exact correspondence, since this nonlinearity is established not at the output of this channel of the regulator, but at its input. Thus, in our case, the width of the dead zone is 0.3; the step of discreteness in the model of the second control channel is 0.2. For comparison, we use two variants of the cost function. In the first variant, this is the integral of the error module multiplied by the time from the beginning of the transient process. In the second variant, we introduce the weighted sum of the squares of the control signals formed at the outputs of the regulator under the integral. Fig. 10.1 shows the block diagram for modeling and optimization of the regulator in accordance with the model of the object and the proposed method for solving the problem in the first variant.

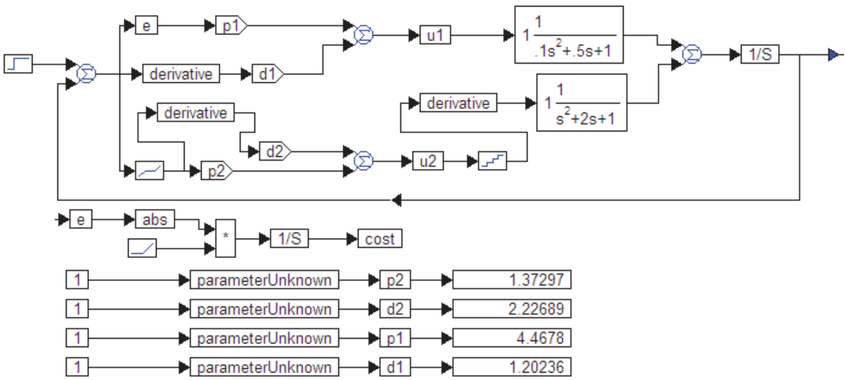


Fig. 10.1. Structure for modeling and optimization of the regulator in software *VisSim*

The starting values of all regulator parameters for the optimization procedure are taken equal to unity. As a result of the procedure, the values are obtained, which are shown in the indicators at the bottom on the right in Fig. 10.1.

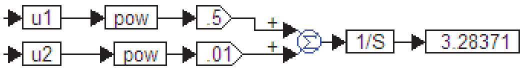


Fig. 10.2. Block diagram for calculating the cost of control resource

Fig. 10.2 shows a block for calculating the resource consumption during the transient process. This consumption was 3.28 units. Fig. 10.3 shows the transient process in the system with the calculated regulator. The duration of the process is little over 6 s, there is a slight overshooting of about 2%, the number of noticeable vibrations is three. The form of control signals is shown in Fig. 10.4.

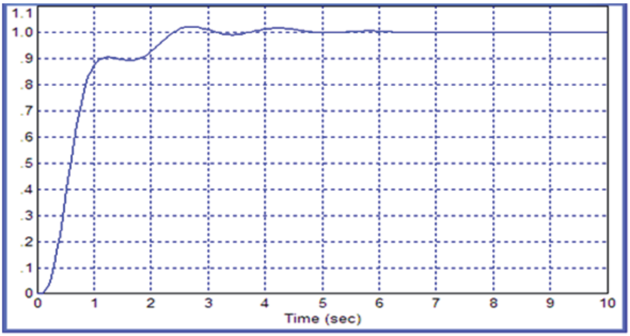


Fig. 10.3. The transient process in the system of Fig. 10.1

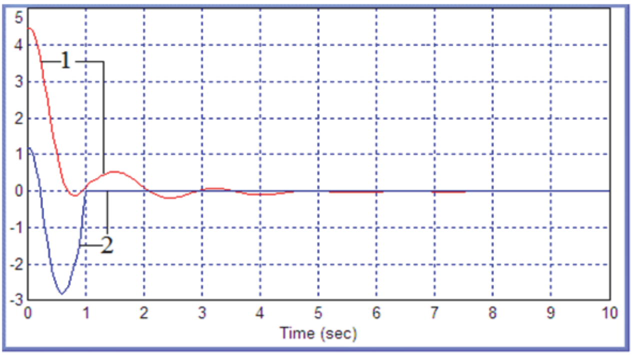


Fig. 10.4. Changes in control signals in two different channels in the system according to Fig. 10.1: line 1 – signal u_1, line 2 – signal u_2

Example 33. Let us enter into the cost function the output of the block for calculating the cost of the control resource and we will repeat the realization of the numerical optimization of the regulator coefficients. The received coefficients of the regulator are also shown in the indicators on the lower right. It can be seen that the control resource cost was 0.608 units, which is more than five times lower than without using this method. The corresponding transient processes in this system are shown in Fig. 10.6 and Fig. 10.7. It can be seen that the transient process is objectively improved, namely: the process time is reduced to 4.5 seconds, that is, a quarter from the previous result. There is no overshooting completely. There is practically no oscillation near the steady state, there is only one wave of backward oscillation, that is, one can say half the oscillation. Reduction of the cost of the control resource is ensured by the fact that the expensive resource of the first control channel is used less: the peak value is reduced from 4.5 units to less than 1.5 units. A cheaper control resource in the second channel is also used more economically: its peak values were +1 and -3, while the steel values were +0.3 and -0.1.

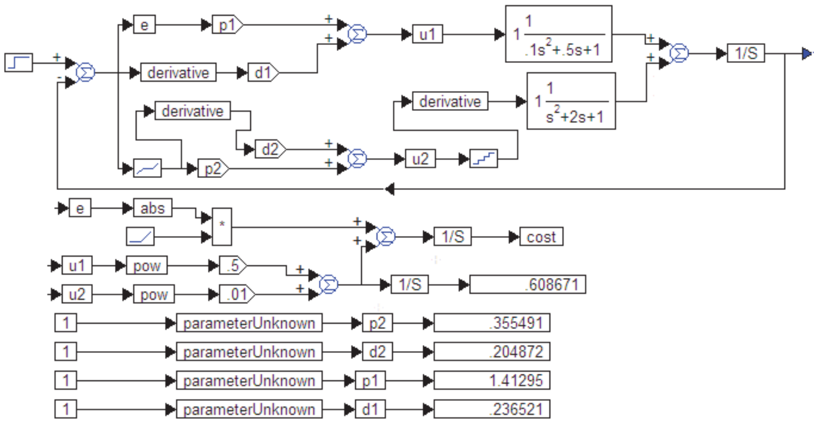


Fig. 10.5. Scheme for modeling and optimization of the regulator taking into account the cost of the control resource and the results of numerical optimization of the regulator coefficients

Note. In the structure shown in Fig. 10.5 the integrator from the resource cost calculator is not used, since the cost function calculator already has its own integrator. It was possible to transfer the adder by setting it at the integrator output, it is not important, saving the number of blocks in the circuit would not be in any case.

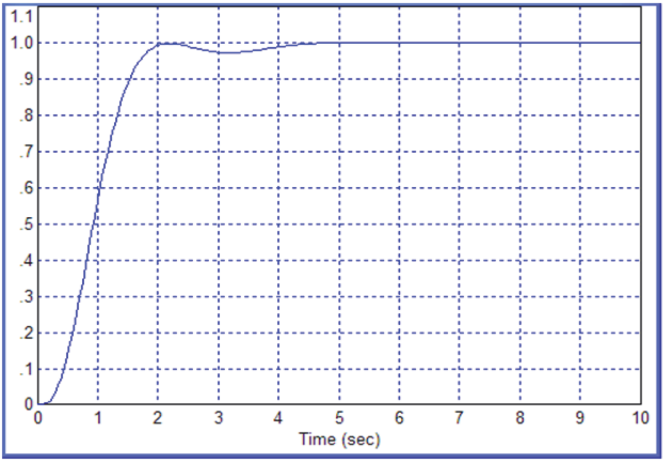


Fig. 10.6. The transient process in the system of *Fig*. 10.5

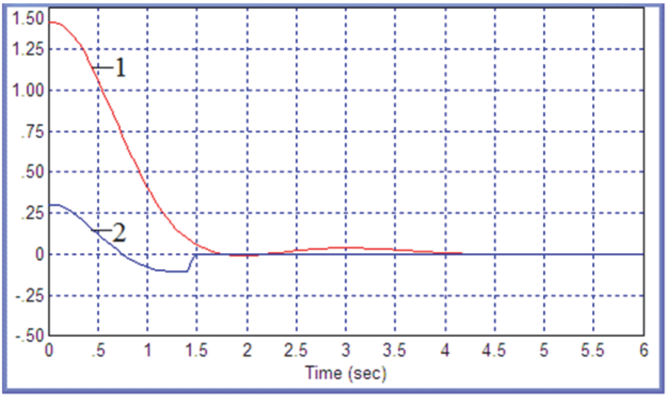


Fig. 10.7. Changes in control signals in two different channels in the system in Fig. 10.5: line 1 – signal u_1, line 2 – signal u_2

Conclusion 13. The introduction to the cost function of the result of calculating the control resource in the form of an integral from the weighted sum of squares allows reducing of the resource consumption, and also improving of the quality of control. This improvement in quality includes increasing the speed, reducing or eliminating the overshooting and reducing the number of oscillations in the transient process.

This conclusion is valid, in particular, with reference to the control system of one output quantity over two channels of influence, including for different resource costs through different channels. This applies, among other things, to the case where one of the channels contains a discrete level converter. All that is said in this derivation is valid for at least one example.

10.2. Robust power saving two-channel regulator with single output object, prone to oscillations

The tendency of the object to oscillate, as a rule, arises from the presence of complex-conjugate roots in the model of the object. Such an object generates a series of oscillations at the output of a stepped action at its input. Feedback using traditional methods of calculating the regulator may be not effective in suppressing these fluctuations sufficiently. In this case, a developer has to use special techniques. In some cases, even a simple increase in simulation time during optimization can be useful. In other cases, more radical measures are necessary, such as creating local or pseudo-local feedbacks, and so on.

Example 34. Let us consider the modification of Example 33, changing the model of the first (better) control channel for the worse, namely: increase the propensity of the object to oscillations. To do this, it suffices to reduce the coefficient in the polynomial of the denominator for s in the first degree, for example, by a factor of five. The new transfer function of the object will look like:

$$W_1(s) = \frac{1}{0{,}1s^2 + 0{,}1s + 1} \cdot \frac{1}{s}. \qquad (10.\ 3)$$

To ensure stability, we increase the simulation interval to 50 s, that is, 2.5 times more. Also we reduce the dead zone by a factor of 1.5 to a value of 0.2. We obtain the following results: $p_1 = 0.646981$; $d_1 = 1.61526$; $p_2 = 0.686261$; $d_2 = 0.118547$. At the same time, the resource cost is $F = 0.28267$.

Optimization in a structure similar to that shown in Fig. 10.1, gives the regulator the following coefficients: $p_1 = 0.467781$; $d_1 = 3.11038$; $p_2 = 1.0038$; $d_2 = 0.956992$. At the same time, the resource cost is $F = 0.467781$.

Fig. 10.8 and Fig. 10.9 show the transients at the output of the system and control in its individual channels. It can be seen that the process in

Fig. 10.8 is unsatisfactory, since oscillations about the steady state develop in it, which not only do not attenuate, but, on the contrary, increase in amplitude, which indicates the instability of the system.

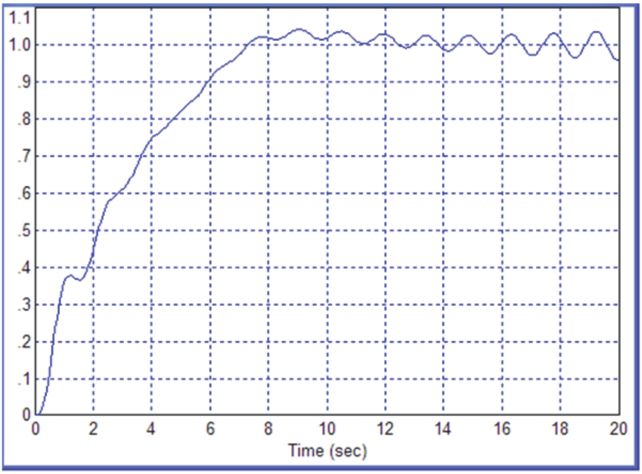


Fig. 10.8. Transient process in the system according to Fig. 10.5 with parameters of the transfer function of the first channel by the relation (10.3)

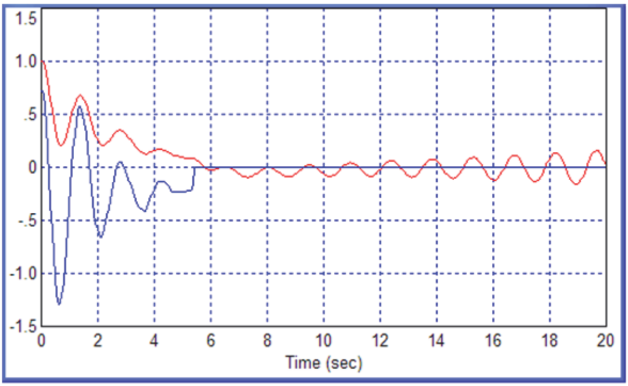


Fig. 10.9. Changes in control signals in two different channels in the system in Fig. 10.5 with parameters of the transfer function of the first channel by the relation (10.3)

Fig. 10.10 shows the transient process for a longer time.

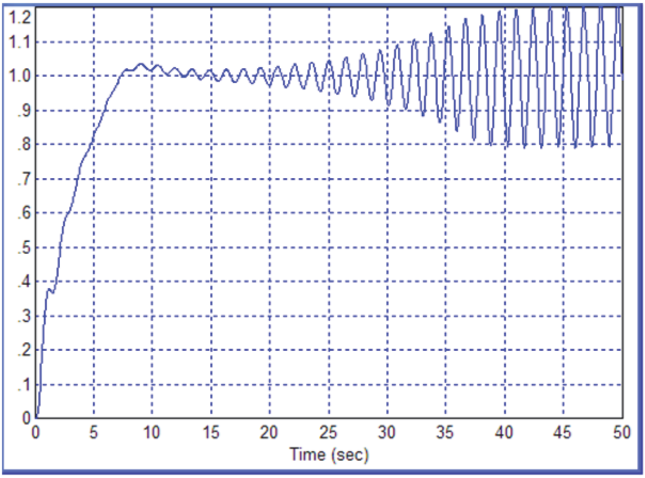


Fig. 10.10. The transient process, under the same conditions as in Fig. 10.8, over a longer time interval

In general, if the simulation interval was changed, the cost of the resource would not be comparable if the integrator was not contained in the mathematical model of the object. In the presence of an integrator, control signals asymptotically approach zero when the transient process is completed, so increasing the simulation duration with stable control does not significantly affect the resource consumption of the control resource, so the results can be compared. Figures 10.11 and 10.12 show the corresponding transient processes in the resulting system. In order to verify with the robust stability of the system, we increase the simulation time by half, that is, up to 100 *s*, and round the calculated coefficients of the regulator so that only two significant digits remain, that is, $p_1 = 0.65$; $d_1 = 1.6$; $p_2 = 0{,}69$; $d_2 = 0.12$.

As it is seen, the process shown in Fig. 10.11 is stable. The difference of this process due to the rounding of the coefficients from the process with non-rounded coefficients is negligible. The graph does not visually show up any difference (the graphs merge). Fig. 10.12 shows the corresponding control signals. The time on this graph is shortened for better visualization and the signal scale is enlarged.

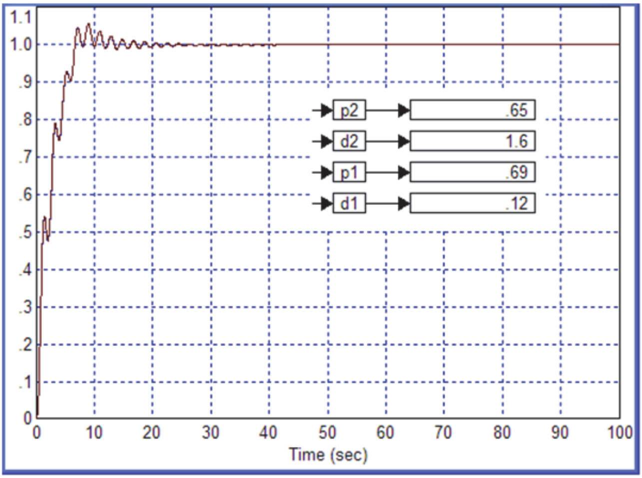


Fig. 10.11. Transition in the system

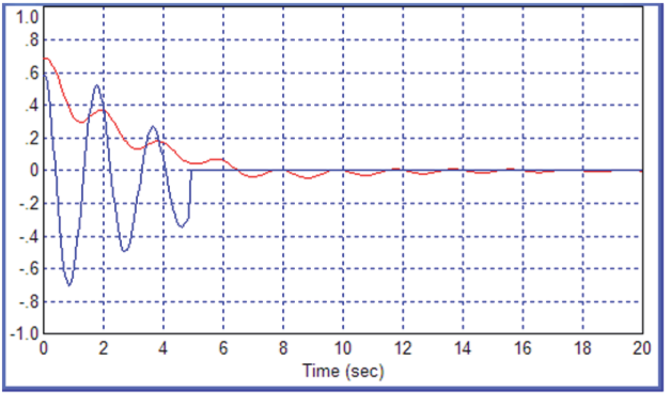


Fig. 10.12. Changes in control signals in two different channels in the system

Conclusion 14. To obtain robust regulators, it may be useful to increase the simulation time.

Conclusion 15. Correct choice of the value of the dead zone can also ensure the success of solving the problem of numerical optimization of the regulator.

10.3. Regulator with separation of "right" and "wrong" movements

In this section, an error growth detector is proposed that identifies those sections of the transient process where the error grows. This is indicated by the positive value of the product of the error, not its derivative. These movements in the system can be conditionally called "incorrect", since ideally the system must perform such movements, as a result of which the error does not increase, but decreases in absolute value.

It has been said above that control of objects that are prone to fluctuations can be extremely difficult. For example, the procedure of numerical optimization may not lead to the finding of regulator coefficients that would provide sufficient speed with a sufficiently high quality of the transient process.

Insufficiently high quality of the transient process can be manifested, for example, in large overshoot, or in a large number of oscillations and, accordingly, their weak attenuation, or in the not monotonic transient process in the initial section.

A solution of the problem is proposed, consisting in dividing the transition process into two types, namely: a) in areas where the error value decreases or is constant; b) on areas where the error value increases. To determine such areas, an appropriate detector is required, which is described above. It consists of a derivative element, a signal multiplier, and a nonlinear element, which is a lower limiter with a level of constraint equal to zero. Such a detector is included in the structure shown in Fig. 9.13. For definiteness, its structure is shown separately in Fig. 10.13. Figure 10.14 shows the structure of the system as a whole.

The proposed system works as follows. In the initial state, the switch connects to its output, and hence to the input of the object, one of its inputs, that is, the output of one of the regulators, the first or the second. Thus, the control loop of the system is obtained closed using one of two controllers. This loop works as in any system with feedback, namely: the output signal of the prescription is subtracted from the input signal of the system, the difference obtained at the output of the subtractor is the error signal $e(t)$. This error signal is converted by one of the regulators into a control signal that is passed through the switch to the input of the object and acts on it to change its output signal to the desired direction. Due to the feedback action, the output signal of the object becomes equal to the prescribed value entering the input of the system. In this case, the detector analyzes the error signal from the output of the subtractor and on its basis generates a logic signal that controls the operation of the switch. Depending on this signal, the output of this switch receives a signal from its first or from its second input. The detector, depending on whether error decreases in size or increase, connects the first or second regulator. Both these regulators are pre-configured by numerical optimization in the circuit that implements such a switch.

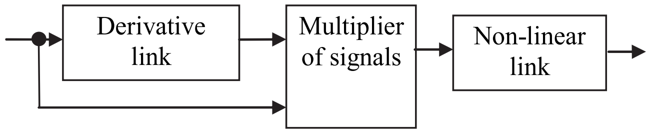


Fig. 10.13. The error growth detector, the same that detector of "wrong movements"

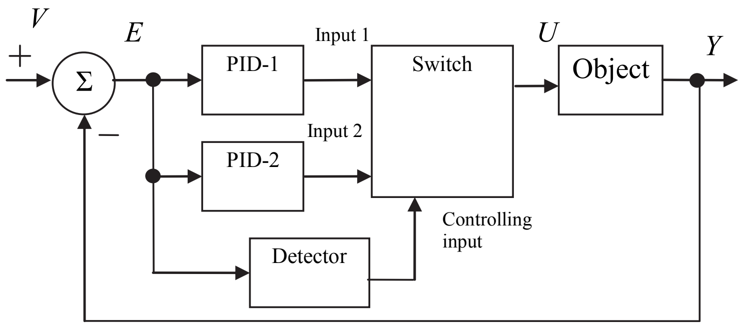


Fig. 10.14. The proposed system with separation of "wrong" and "right" movements: PID is PID regulators

The theoretical basis of the method can be given on the basis of the following considerations. Since the "right" and "wrong" changes in the output signal of an object in the system can alternate, it can be interpreted as alternating "fight" and "wrong" operation of the regulator. Therefore, the question of correcting the "wrong" operation of the regulator by changing its coefficients may be raised. To test the productivity of this idea, it is sufficient to simulate such a system, while both regulators can have the same mathematical models, but different gain factors, which are determined by the method of numerical optimization. If this method is not effective, the numerical optimization procedure should give the same coefficients for both regulators, since switching regulators does not lead to a decrease in the value of the cost function. If the simulation shows that the procedure always gives different coefficients for the two regulators, then this can be considered a confirmation of the effectiveness of the method for the types of regulators studied.

This system can be further improved, as shown in Fig. 10.15. The positive effect of the improvement is that there is no need to switch the integral link of the regulator. Therefore, it is proposed to include this regulator link in addition to the switch, directly to the input of the object, but an adder for these purposes is necessary.

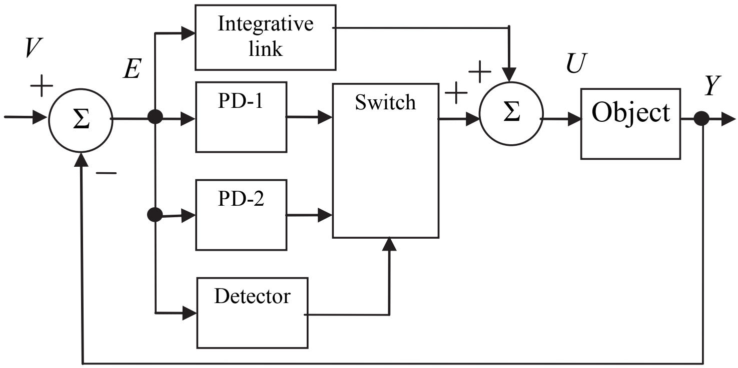


Fig. 10.15. An improved version of the system in Fig. 10.14

The system proposed in Fig. 10.15 takes into account the internal structure. It is much simpler than the system shown in Fig. 10.14, since in the system in Fig. 10.14 in each of the regulators there are three links, and one combiner with three inputs, and in the system of Fig. 10.15, in each of these regulators there are only two links and one adder with two inputs. In the system of Fig. 10.14 there are six configurable parameters, and in the system of Fig. 10.15 there are only five such parameters (two derivative links, two proportional links and one integral link). Simplification is not the goal as it is: the traditional PID regulator is even simpler, but in the structure in Fig. 10.15 there is positive effect that justifies this complication.

Example 35. To illustrate the effectiveness of the proposed method, Fig. 10.15 shows a simulation of such a system in terms of the above structure. In this case, the mathematical model of the object is given in the form of a transfer function of the following form:

$$W_O(s) = \frac{1}{s^3 + s^2 + s + 1}. \qquad (10.4)$$

Fig. 10.16 shows a project for modeling in VisSim of a system according to the structure shown in Fig. 10.15. In this structure, three composite blocks are used: the regulator (named "PI-regulator"), "Optimizer" and "Cost Estimator". When modeling, different types of regulators were used, not only PI-regulator, but the block name was saved so that the entire project could not be altered, but only by editing the block structure. Fig. 10.17 shows the internal structure of the regulator; Fig. 10.18 gives the internal structure of Cost Estimator, and Fig. 10.19 presents structure of Optimizer.

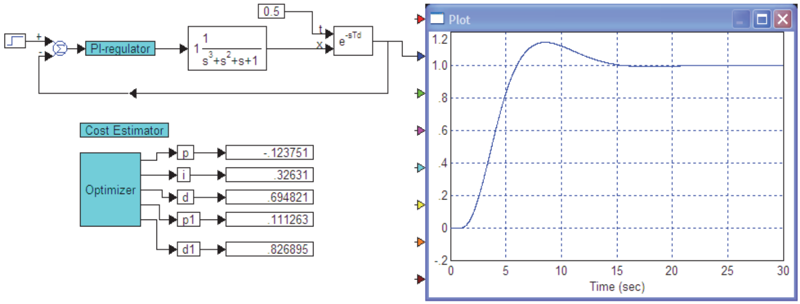


Fig. 10.16. Structure and transients processes of Example 35

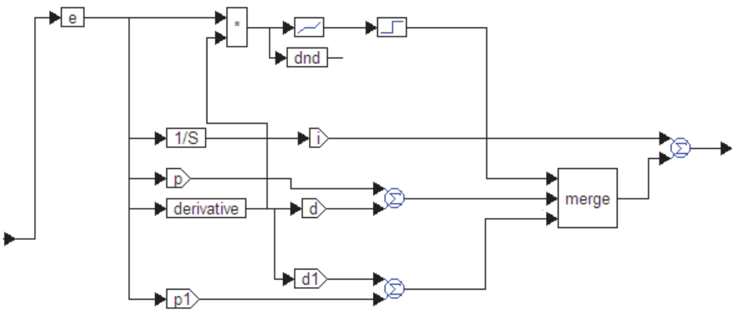


Fig. 10.17. The structure of the regulator PI-regulator
(here PID regulator is used)

In Cost Estimator, there is a calculator of the integral of the error module multiplied by the time. Also the output signal of the detector of incorrect movements is introduced under the integral with the factor equal to ten. This signal is indicated by the variable "dnd". As can be seen from Fig. 10.17, in this detector, the control error indicated by the variable e is multiplied by the derivative of the error by the multiplier indicated by [*]. The derivative is calculated by the unit "derivative".

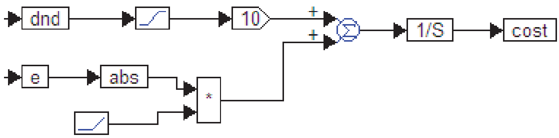


Fig. 10.18. The structure of the Cost Estimator

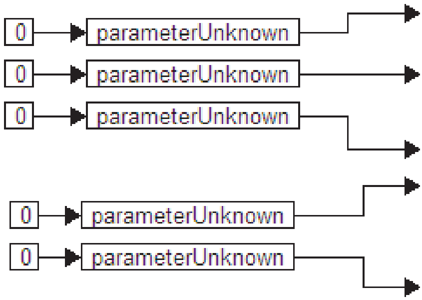


Fig. 10.19. Structure of the block Optimizer

The product of the error on its derivative from the output of the multiplication unit is fed to two serially connected nonlinear elements: a limiter and a relay link. Together they form the required nonlinear element, shown in Fig. 10.13. If the error increases in magnitude, then the product is positive, and if it decreases in magnitude, then this product is negative. A nonlinear element converts this signal into a discrete signal with two output values. This signal goes to the input of the merge block (switch), which connects one of its analog inputs to its output, that is, realizes the functions of the switch. The coefficients are calculated in the optimization mode provided in the *VisSim*. For comparison, a simple PID regulator was used. The graph obtained with the indicated system is shown in Fig. 10.16 on the right, and the graph obtained with a simple PID *regulator* is shown in Fig. 10.20 on the right. It can be seen that the quality of control in Fig. 10.16 is better than the control quality in Fig. 10.20. Indeed, in the first case, the process is initially monotonic, until it crosses the level of the prescribed value. Then there is an overshooting of about 15%, after which the process smoothly (asymptotically) moves to the prescribed value with out additional oscillations. In the second case, the process in the initial stage goes in an erroneous direction, not upwards, but downwards, that is, there is an inverse overshooting. Then it moves in correct direction, but after reaching the prescribed value, it has a small overshooting, but then again deviates, and the overshoot reaches the same value of 15%. There are oscillations, at least, with four clearly distinguishable maxima.

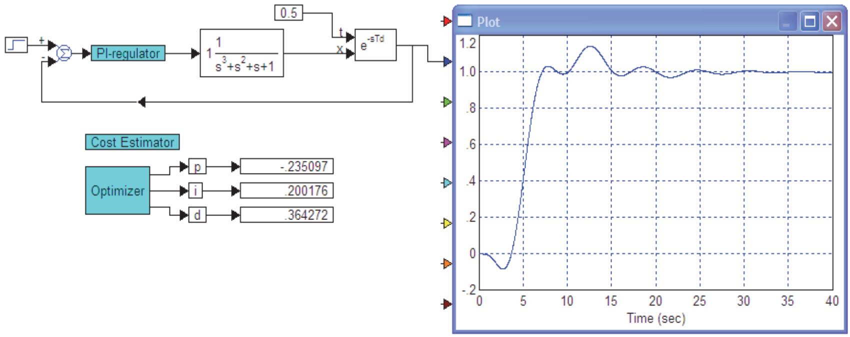


Fig. 10.20. Structure with a simple PID regulator and transient process

Reverse overshooting refers to the significant shortcomings of the system. Indeed, only in exceptional cases it is possible to reconcile with it. It is sufficient to imagine that in the case when it is necessary, for example, to cool an object, the system first heats it, and then cools down to the desired temperature. Or, let's say, when a maneuver is required to turn left, the system first performs a small turn to the right, and only then performs the required turn to the left. Therefore, the quality of the system in Fig. 10.16 is much higher than the quality of the system in Fig. 10.20.

Also an indirect sign of the effectiveness of the proposed method is the fact that the optimization procedure gave essentially different values for the PI regulator coefficients, they even have different signs.

These values are indicated in the output displays of the resulting variables. Namely: in the case of applying a simple PID regulator, the calculation gave the following regulator coefficients: $K_p = –0.235$; $K_i = 0.200$; $K_d = 0.364$ are, respectively, the coefficients of the proportional, integrating and derivative links. In the case of applying the system of Fig. 10.15 and 10.16, the coefficients for the first regulator are obtained: $K_{p1} = –0.123$; $K_{d1} = 0.695$; coefficients for the second regulator: $K_{p2} = 0.111$; $K_{d2} = 0.286$; Coefficients for the third regulator $K_i = 0.326$.

10.4. The use of bypass channel for feedback control of an oscillating object

The complexity of solving the control problem is determined by the complexity of the model of the object. If the model of the object is such that even with small input signals at the output of the object several oscillations of different frequencies with increasing amplitudes are formed, then the control of such an object can turn out to be quite complicated.

The models of some objects may differ insignificantly in appearance of the mathematical record, but the calculation of the regulator for some of such objects can be extremely simple, and for others it is extremely difficult.

In this chapter, it is proposed to use a simple PID regulator of the form (2.2) or (2.3), supplemented by a bypass channel [23], which is connected in parallel to the control object, as shown in Fig. 10.21. This structure is similar in structure to that of Smith's predictor [45–48], but the mathematical model of the bypass channel is calculated differently, so the bypass channel is more flexible. *Smith* predictor can be considered as a special case of a broader approach called "bypass channel".

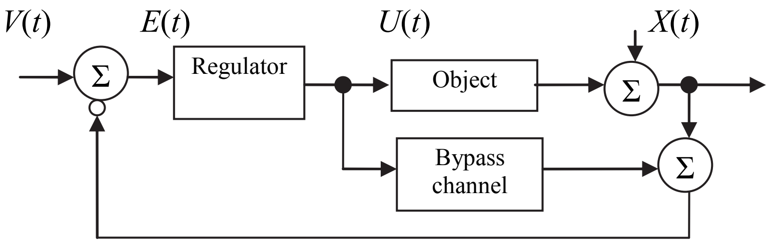


Fig. 10.21. Structure of the system with bypass channel

The *Smith* predictor can be implemented only if the object contains a link of pure delay, for example, the transfer function of an object is equal to the product of two functions:

$$W(s) = W_1(s) W_2(s) . \tag{10.5}$$

In this case, the first factor is defined, for example, by the relation:

$$W_1(s) = \frac{b_0 + b_1 s + ... + b_m s^m}{a_0 + a_1 s + ... + a_n s^n} . \tag{10.6}$$

Here m and n are the order of the numerator and denominator, and the second factor has the following form:

$$W_2(s) = \exp(-s\tau)\,. \tag{10.7}$$

The *Smith* predictor equation in this case has the following form [23]:

$$W_{PS}(s) = W_1(s)[1 - W_2(s)]\,. \tag{10.8}$$

It is easy to see that the total transfer function of the two channels, that is, the channel of the object and the channel of the *Smith* predictor, is in sum equal to the transfer function of the form (10.6) and is free from the factor of the form (10.7). Indeed,

$$W(s) + W_{PS}(s) = W_1(s)\,. \tag{10.9}$$

We can assume that *Smith* predictor allows compensating for the effect of delay on the operation of the stabilization loop. Nevertheless, the situation is not as good as follows from (10.9). First, the output of the adder, which sums the output signal of the object with the output of the *Smith* predictor, is still not the output signal of the object (and system), so the stability of the system as a whole does not guarantee a qualitative transient processes. This drawback also applies to the system using the bypass channel, which is taken into account in further studies. Secondly, if there is no delay in the object model, then $W_2(s) = 1$, which implies $W_{\mathrm{PS}}(s) = 0$, that is, *Smith* predictor cannot be applied.

The bypass channel proposed in [23] is calculated on the basis of the following principles:

1. The low-frequency part of the frequency response of this channel must be much lower than the low-frequency part of the frequency characteristic of the object.
2. The high-frequency part of the frequency response of this channel must be much greater than the high-frequency part of the frequency characteristic of the object, and at the same time it should be such that the closed loop would be stable with the best quality.

This method is developed for the case when the amplitude-frequency characteristic of the object at certain frequencies decreases sharply in magnitude and the phase-frequency characteristic, starting from certain frequencies, increases sharply in magnitude. In this case, ensuring sustainability is difficult. The bypass channel, having a large amplitude-frequency response in the high-frequency region, is crucial for stability, since in comparison with the output signal of this channel the output signal of the object in the high-frequency region can be neglected. This makes it possible to ensure the stability of the control loop, even when the frequency characteristic of the object is most inadequate to meet the requirements for it to ensure stability and quality of control. Thus, the loop remains stable. At the same time, since the amplitude-frequency characteristic of the bypass channel is negligibly small in comparison with the amplitude-frequency characteristic of the object in the low-frequency region, the presence of the bypass channel does not affect in this region. The signal at the output of the adder in this frequency range is mainly determined by the output signal of the object.

The newest concept of the bypass channel extends these requirements to the next set:

1. Requirement 1 is retained completely.

2. Requirement 2 is modified to the following requirement: "The transfer function of the bypass channel is such that its sum with the transfer function of the object gives an object whose control in the loop with negative feedback is successfully provided by any of the methods, for example, by numerical optimization of the regulator".

This modification of this requirement is broader. In the initial concept, the bypass channel completes only the high-frequency part of the object's transfer function to such a transfer function, which is more convenient for designing of an effective regulator. In the new concept, the bypass channel can contain any terms that, in conjunction with the transfer function of the object, will give a new transfer function more convenient for control, and these terms do not necessarily only refer to the high-frequency part of the object model.

Thus, in this chapter, a modified approach to designing a system using a bypass channel is proposed.

In particular, one of the possible ways to calculate the transfer function of the bypass channel can be the following algorithm.

1. First, the desired transfer function $W_D(s)$ is sought, that is, the function closest to the transfer function of the object, but more convenient for designing the regulator.

2. Subtraction of the transfer function of the object from the obtained desired transfer function results in a preliminary transfer function of the bypass channel.

$$W_{BC}(s) = W_D(s) - W(s). \tag{10.10}$$

3. If the obtained transfer function is not very convenient for implementation, then its final form can be obtained by some simplification, with the following rule: if in some frequency range the transfer function obtained in the point 2 is much less than the transfer function of the object (at 30 and more times), then in this frequency range it can be changed arbitrarily, provided this property is preserved, that is, that this function remains much less than the transfer function of the object (by 30 or more times).

Example 36. Let us consider an object whose transfer function has the form of a rational fraction in the region of Laplace transforms:

$$W_1(s) = \frac{s^3 + 4s^2 - s + 1}{s^5 + 2s^4 + 32s^3 + 14s^2 - 4s + 50}. \quad (10.11)$$

If all the coefficients of the polynomials in the numerator and denominator were positive, the object could be quite simple to control. But even in this case, the combination of coefficients can be such that the control of such an object becomes more complicated. If among the coefficients there are zero and even negative coefficients, as in (10.11), then the design of the regulator can be extremely difficult. Several publications investigate the possibility of solving the object control problem (10.11), and its transfer function is called complex. The complexity of the object is that when the regulator is numerically optimized, the result is a system prone to reverse overshoot. But direct overshoot is also great. Both these kinds of overshoot are highly undesirable. In work [43], the struggle with both types of overshoot is accomplished by increasing the order of the regulator, where along with the first derivative, the second, third and even fourth derivative of the error is used. This greatly complicates the regulator, and in addition, as known, derivation sharply emphasizes noise, as a result of which it can be implemented only in a strictly limited frequency band. This method is also not effective enough, although it is very difficult. Preliminary investigations have revealed that the most unfavorable factor in the model of an object of the form (10.11) is the negative sign of the coefficient of the polynomial in the numerator. If the sign before the term in the first degree is replaced by the opposite one, then the design of the regulator by the method of numerical optimization gives a satisfactory result. It follows that the fractional function, whose denominator coincides with the denominator of the transfer function of the object, can serve as the transfer function of the bypass channel, and the numerator is "+ 2*s*". The absence of a free term in the numerator automatically leads to satisfaction of the first condition, since in the low-frequency region the value of such a transfer function asymptotically tends to zero.

Example 37. We will calculate the high-order regulator for the object from Example 36 based on this principle with the bypass channel. Let the regulator along with the proportional and integral links include link of higher derivatives, from the first through the fourth inclusive. The corresponding block diagram for modeling and optimization in the *VisSim* is shown in Fig. 10.22. The resulting transient is shown in Fig. 10.23.

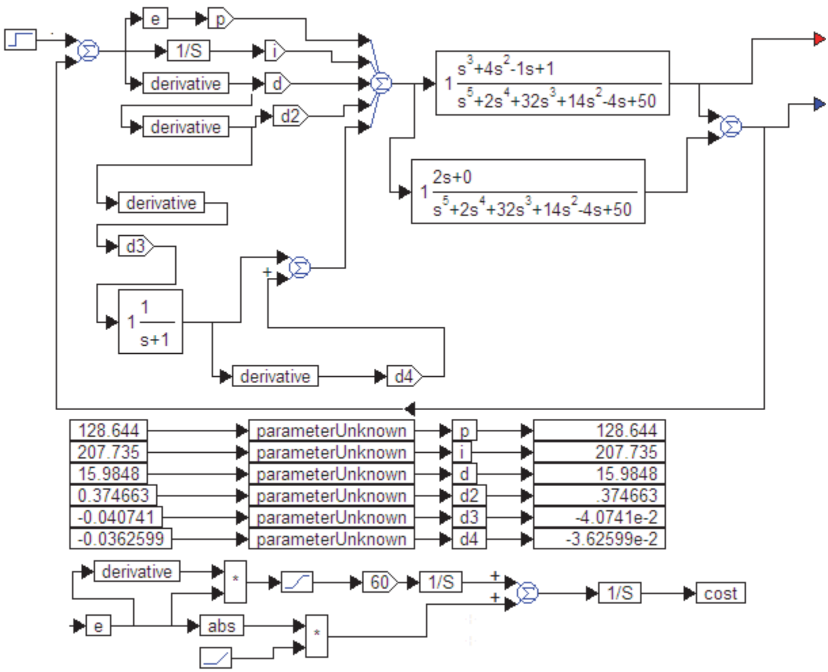


Fig. 10.22. Block diagram of the system with PID^4-regulator

It is important to note that the process at the output of the adder, which is sums the output signal of the object and the output of the bypass channel, is not essential for assessing the quality of the system. This is just the sum of the real and virtual signals. Only the output signal of the object is important. Therefore, although the system successfully controls the composite object that contains bypass channel together with the real object, the result should be evaluated by the transient process at the output of the object. From the analysis of the signal in Fig. 10.23, it is seen that the output signal of the composite object is characterized by high quality. Namely: overshoot does not exceed 10%, the difference of this signal from a single jump is unimportant, that is, the error is almost completely zero after the twelfth second from the beginning of the transient process: in this case, there is no reverse overshooting, and the reverse deviation from the prescribed value after the first positive overshooting also does not exceed 10%. Nevertheless, the transient process at the output of the real object is much less attractive. Namely, overshooting is 40%, there is no reverse overshooting, but the reverse deviation is about 85%. The second reverse deviation is slightly less than 20%, the second overshooting is slightly less than 10%.

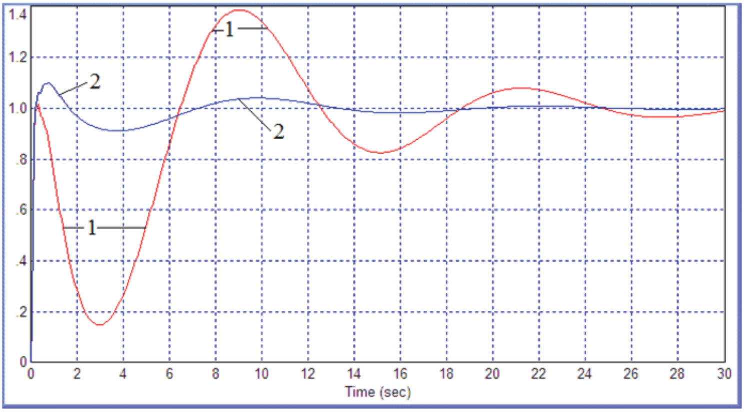


Fig. 10.23. Transient processes in the system according to Fig. 10.22: line 2 is the total output, line 1is object output

The duration of the transient process is more than 30 *s*. Still, the advantage of this method is that the result is obtained quite easily, the reverse overshooting is not complete (we can talk about a margin of about 15%, that is, the output value does not reach its starting value by 15%, which allows speaking about a reliable absence reverse overshooting). The disadvantages of this solution include the excessive complexity of the regulator, since in this case, together with a fourth-order PID regulator (PID^4 regulator), an additional bypass channel of the fifth order is used.

Example 38. Let us consider the solution of the same problem in the same way when using a simple PID regulator. In doing so, we do not use the virtual error $e(t)$ signal, which is calculated by subtracting the output signal of the composite object from the signal, and the actual error signal $e_1(t)$, which is specially calculated for these purposes by subtracting the actual output signal of the object. The structural diagram for this case is not shown, since all modifications of it are understandable from the text. The resulting transient is shown in Fig. 10.24. It can be seen that as a result, the signal at the output of the composite object, like the signal at the output of the actual object, has not changed in the main parameters. In this case, the modeling time is increased to show clearly, then after 40 *s* the both transient processes practically merge, the control error is close to zero, since the output signal of the object is close to unity.

When using intermediate types of regulators, namely, PID^2 and PID^3, the results approximately coincide. With respect to the result with the simplest variant, that is, with a PID regulator, it can be concluded that the received regulator is practically inferior in its parameters to all the regulators considered above, although it is much simpler and does not require even a second order of derivation. This regulator is much easier to calculate, the procedure for its numerical optimization is almost self-evident with taking into account the given the idea of its design, which is outlined in the three theses above.

For comparison, Fig. 10.25 shows the transient process obtained in [44] for a system with the same object.

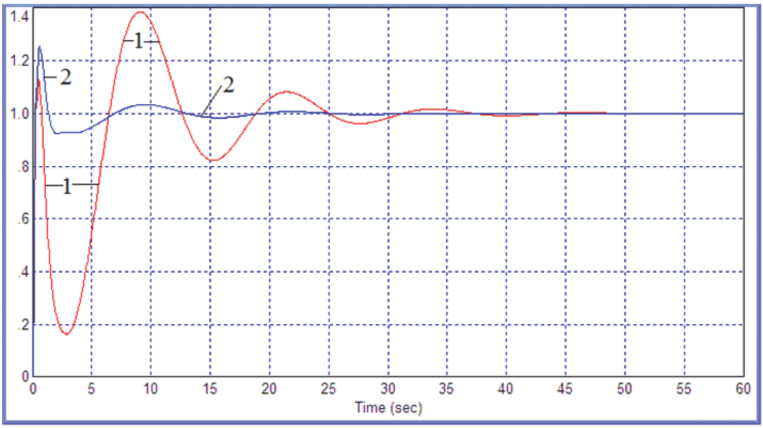


Fig. 10.24. Transient processes in the system from Example 38: line 2 is total output, line 1is object output

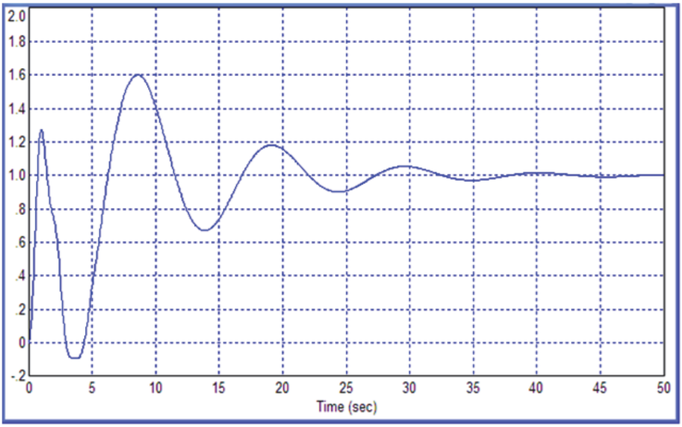


Fig. 10.25. Transient processes in a system with the same object in accordance with [43]

This process is clearly worse, since overshooting is 60% (in our result it is only 40%), the duration of the process is the same in both cases, the reverse overshoot is 10% (in our result it is absent). In our case, the reverse movement slightly exceeds 80%, whereas in this case the reverse overshooting in the process shown in Fig. 10.25 is 110%. More than 100% gives backward overshooting, less than 100% is only backward movement.

11. FEEDBACK SYSTEMS WITH PSEUDO LOCAL LOOPS

11.1. About local and pseudo local loops

This chapter proposes a method for designing of regulators for unstable problematic objects. The method basis is in introducing of compensative feedback directly into the model of object. In the result of this the structure should be transformed to get single-loop regulators. The method has been investigated with modeling in the software package *VisSim*, and its effectiveness has been demonstrated before. The effectiveness of regulators for some kinds of objects, for example, such as objects with deep non-linear positive feedback, depends on the accuracy of the realization of compensative elements.

The design of regulators for objects containing problematic circuits is a difficult task, especially in the presence of non-linear elements in the case of positive feedback in the individual elements of the structure of an object. Positive feedbacks are very complicated when the loop gain is greater than unit.

To solve this problem the modeling method is proposed and examined. It consists of two stages.

In the first stage compensating feedback loops are introduced into the model of the object. These loops allow to convert the non-linear element into the linear one, or to make its model close to a linear one. Also, these loops can convert integrator into aperiodic link. The resulting modified object is easy for controlling. So the design of the regulator for it is carried out simply on the basis of automatics control theory.

In the second stage the feedbacks loops introduced to the structure must be recalculated into corresponding elements of the successive regulator. In this case, the loops inside the regulator itself are not a problem; they can be preserved in the model of the regulator. The problem is only introduced loops in the structure of object, because they can't be realized directly. So these loops should be recalculated into the equivalent fragments of the successive regulator structure.

Despite the fact that this method is intuitive and clear, with simulating desired positive effect is achieved not always. Simulation with the use of program *VisSim* has the advantage that the calculation of derivatives and integrals is then carried out on the same or similar algorithms, which should take place in the digital regulator. This allows to identify possible problems of implementation, and to find an effective solution to overcome them.

11.2. The control of objects with two integrators

Theoretically, the control of an object, which is a two series-connected integrators, is not of great complexity. But this example can demonstrate the essence of the method and its efficiency.

Let the model of the object is given by the following transfer function:

$$W(s) = \frac{X(s)}{U(s)} = \frac{1}{s^2}. \qquad (11.1)$$

This object can be represented as two serially connected first-order integrators:

$$W(s) = \frac{1}{s} \cdot \frac{1}{s}. \qquad (11.2)$$

If we introduce a negative feedback around one of the integrators, for example, proportional link with the gain equal to two, it will be converted to an aperiodic link. Remaining integrator can play role of an integral regulator for astatic control, i.e. it will reduce the static error to zero.

Thus, to control the object in question it would be sufficient to use the specified local negative feedback, serial regulator with the unit gain and a single global negative feedback, as shown in Fig. 11.1. In this figure, the traditional notation is used for input, output, and control signals. Any element which does not belong to the object is a part of regulator.

In the implementation of this structure a problem appears: the signal $z(t)$, required for switching of local feedback is unavailable for measurements. Simulation of an object, namely – the regulator, provides a model of the signal as the signal $z'(t)$, as shown in Fig. 11.2.

Fig. 11.3 shows the structure for simulation of the system accordingly with the structure shown at Fig. 11.2, with the use of the program *VisSim*. Fig. 11.4 shows the resulting graph of the transition process in response to a single job step jump $v(t)$.

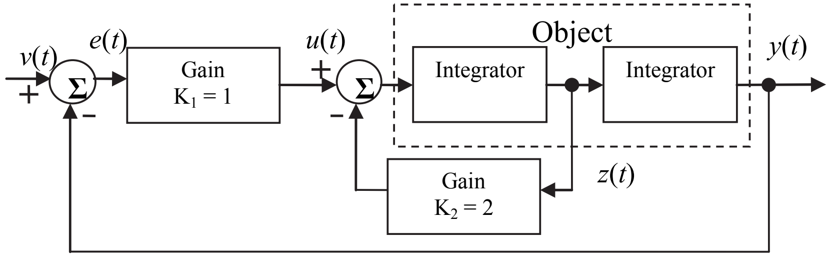


Fig. 11.1. Block diagram for the control of the object with the help of the local loop

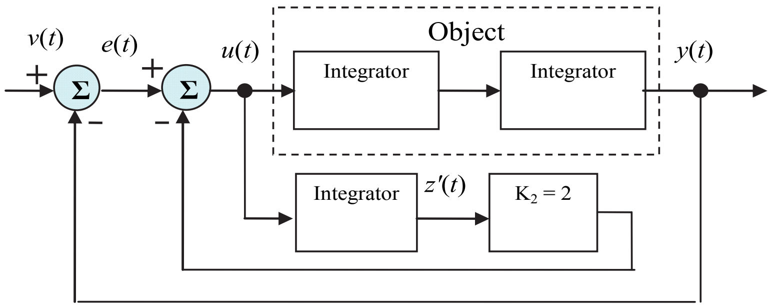


Fig. 11.2. Block diagram for object control using the pseudo local loop

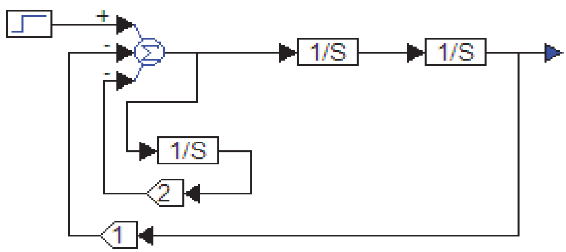


Fig. 11.3. The structure for simulation of the system accordingly with the structure shown at Fig. 11.2: stabilization of object, consisting of two series-connected integrators: a first integrator is stabilized with the help of pseudo local loop, a second integrator of object is not stabilized, and it performs the function of an integral regulator

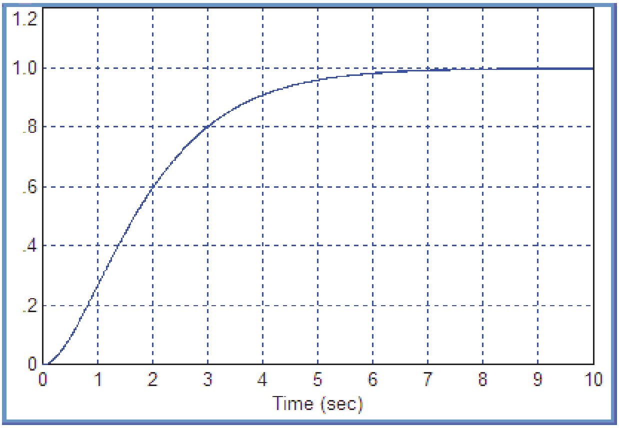


Fig. 11.4. Transient in the system according Fig. 11.3, which shows the absence of overshoot, a static error is zero because of the presence integrator circuit

In the example shown the robust control is achieved, because the object is linear. The exact values of the coefficients of regulator do not matter much. The proposed method is efficient, intuitive and easy to implement in the case of a linear object with small amount of a series-connected integrators without feedback (in our example object is the two integrators).

11.3. The control of objects with two integrators and non-linear positive feedback

Let complicate the problem by introducing a positive feedback, which contains a non-linear element in the form of an element that is raising the input signal into the third power. The structure of such an object is shown in Fig. 11.5. The effective regulator is required to be found. Introduced positive feedback destroys the stability of the object to a great extent. If the output signal is greater than unit, then in the local loop of the object the positive feedback greater than unit arises. The cubing increases the problem. Methods of numerical optimization of PID-regulator cannot be used to calculate regulator for the control of such an object.

If we virtually introduce the negative feedback into the object, repeating the existing non-linear relationship, but with an opposite sign, then these two loops should compensate each other. As a result, we obtain an object that was discussed in the previous section. The controlling of this derived object can be accomplished by the previously demonstrated way. Fig. 11.6 shows the corresponding block diagram.

After that the local non-linear loop must also be converted into a pseudo-local loop to provide a regulator structure, which does not use the signals from the inside of the object and does not use non-existent inputs of the object. Block diagram obtained by means of the necessary equivalent transformations is shown in Fig. 11.7 Simulation fully confirms the effectiveness of the proposed method to the object in question. Simulation scheme is shown in Fig. 11.8 and the results for different values of the task $v(t)$ is shown in Fig. 11.9.

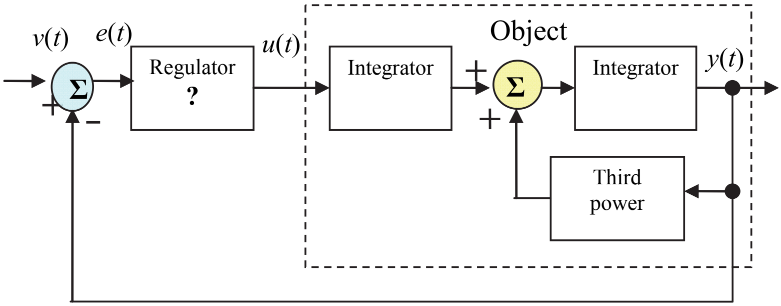


Fig. 11.5. The task of controlling the object with an internal nonlinear unstable loop

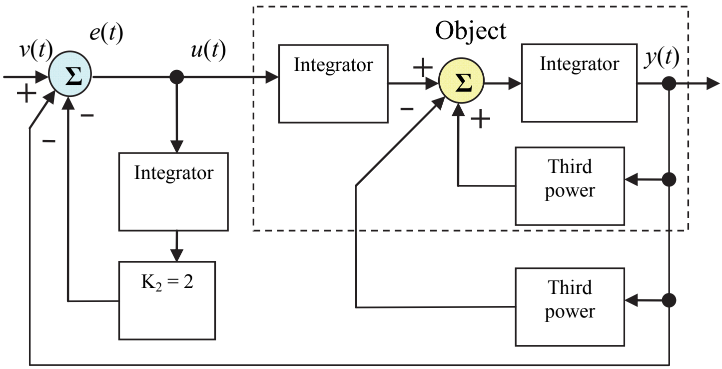


Fig. 11.6. The block diagram for controlling the object using one local and one pseudo local loop

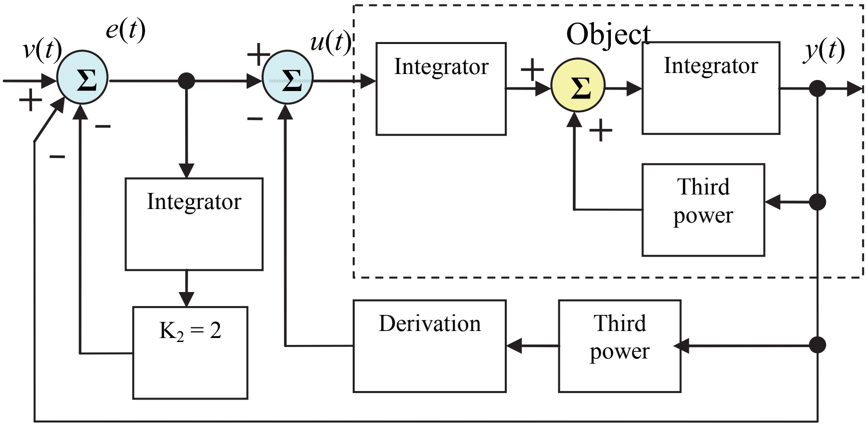


Fig. 11.7. Block diagram for controlling the object using the two pseudo-local loops

The resulting structure of the regulator can also be simplified with the aid of equivalent transformations. When using the digital controller it is not required because this structure can be provided by the according program. The resulting system is fairly robust, that is, small changes in the coefficients do not lead to loss of stability. Also, the stability of the system is not affected by changing the method of integration (i.e. the method of computation of integrals and derivatives of the signals used).

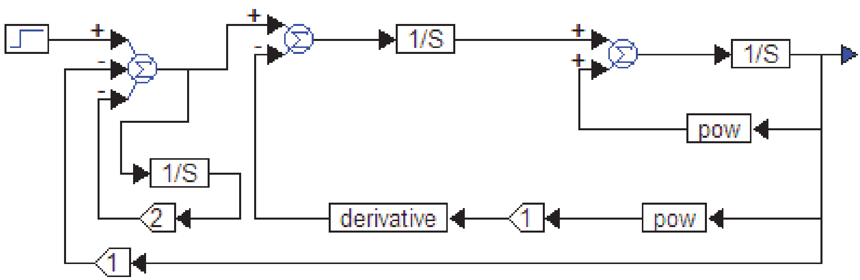


Fig. 11.8. The simulation results of the object: the object is fully stabilized by the pseudo-local feedback

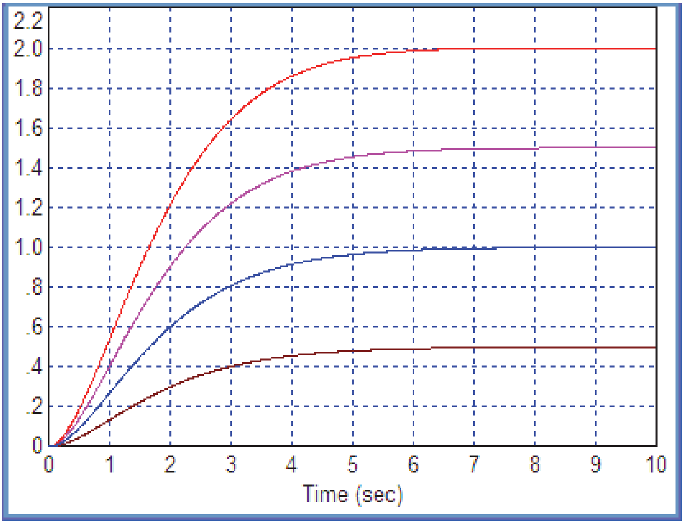


Fig. 11.9. The simulation results of the object for Fig. 11.8: the object is fully stabilized by the pseudo-local feedback; transients is shown for different values of the task (from 2 to 0.5)

11.4. The control of objects with three integrators and non-linear positive feedback

Let us complicate the problem by introducing into the object of another series-connected integrator.

Two variants can be offered for the use of the proposed method.

The first variant assumes deeper negative feedback than inside the object. This makes the problem loop enveloped by the aggregate negative feedback instead of the positive one, as shown in Fig. 11.10. In the simulation of the system stability when input value jump to about unity value has been provided. The nonlinear system may have different levels of quality of transient processes according to the input signal value, for example, it might be stable at low input signals and unstable at high input signals.

Simulation according to the structure shown in Fig. 11.10 reveals a number of problems. We would like to recommend choosing simple Euler method among the possible methods of integrating. This method uses the calculation of the integral of a function with the help of the sum of the values of this function on all the interval of integration, taken with the regular period, multiplied by the duration of such period. Specified above interval is the value of the integration step. Other methods of integration do not give the desired effect.

Fig. 11.11 shows result of the simulation of the converted system, wherein the regulator structure is transformed into a structure with a single main loop (the local control loops inside the regulator are not forbidden). In this structure, naturally, the coefficient *k* in the regulator after involution into the third power can be equal to unity or can be less.

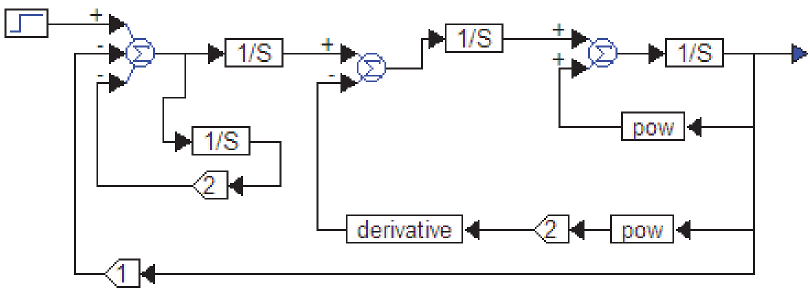


Fig. 11.10. Providing of the stable control by means of the introduction of local loops after the conversion of the primary loop: a first integrator is stabilized by pseudo-local loop, the second block, which consists of two integrators, the latter of which is enveloped by a non-linear positive feedback, is stabilized by a local loop

In the case of a unit value of *k* almost complete compensation of nonlinearity occurs, but the loop has two integrators, making it unstable. Therefore, we use different coefficients *k* greater than unity.

With the gain $k = 2$, according Fig. 11.12, the transient process in response to a single step jump has the best form, as can be seen from the graphs in Fig. 11.13. If the value of this jump is reduced, there is an overshoot in the system. Particularly, when a jump is 0.6, the overshoot amount is about 30%. In the case of increase in this value, in the system arises reverse process, which can be called conditionally "undershoot" as antonym to "overshoot". In particular, when the input signal is equal to the output 1.2 the signal first rather rapidly approaches the value of 1 and then moves slowly to the desired value of 1.2. Fig. 11.14 shows the transient processes depending on the coefficient on the negative feedback *k*.

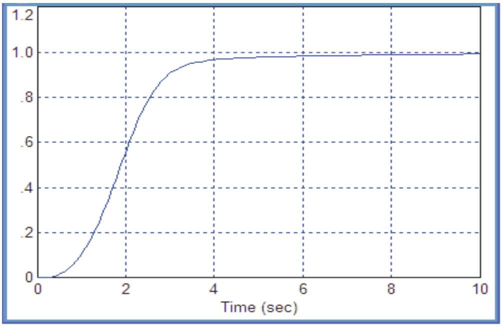


Fig. 11.11. Transient responses for Fig. 11.10

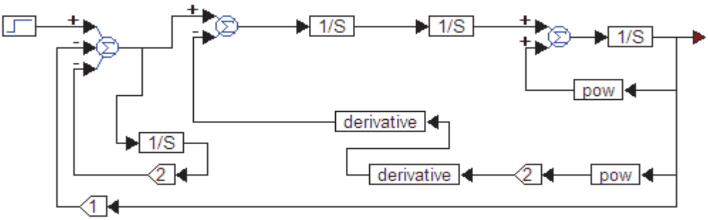


Fig. 11.12. The result of the conversion of the previous structure into the structure with one main loop

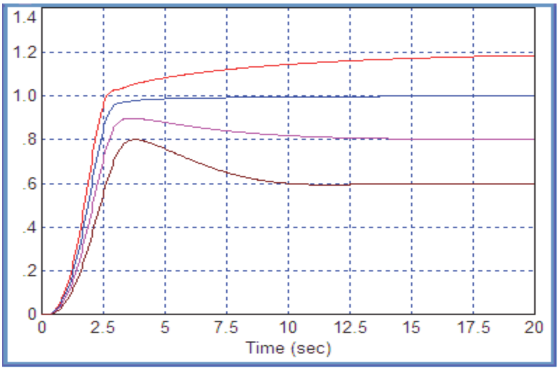


Fig. 11.13. The result of the use of the structure according Fig. 11.12 with one main loop: graphics of the transient processes when at the input step is equal to 1.2; 1.0; 0.8 and 0.6, correspondingly

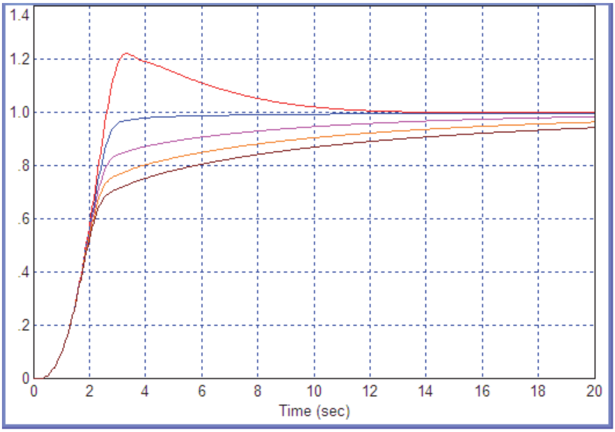


Fig. 11.14. Transient processes in the system with the different gains after the block of cubing: upper graph is with $k = 1.5$, further - $k = 2$; $k = 2.5$; $k = 3$ and $k = 3.5$

As we can see, if the input jump is equal to unity, than for $k = 1.5$ the overshoot exceeds 20%, for $k = 2$ there is no overshoot, when $k = 2.5$ undershoot at about 20% has place in the system, with a further increase of this gain, undershoot slowly rises.

With the reducing of the size of the input signal to a value of 0.5, this dependence varies. Corresponding transients are shown in Fig. 11.14.

When $k = 1.5$ the overshoot exceeds 80%, for $k = 2$, overshoot is 50%, with k = 2.5, it is 30% and with further increasing of gain k the overshoot is reduced.

With the increasing the value of the input signal to a value of 1.2, this dependence is also changed. Corresponding transient processes are shown in Fig. 11.15. When $k = 1.4$ the overshoot is close to 20%, with $k = 1.6$ overshoot is negligible, when $k = 1.8$ or more, undershoot occurs. If $k < 1.2$, then the system is unstable.

The second variant of controlling uses the attempt of a total compensation of the nonlinearity. In this case, discussed above ratio should be equal to unity: $k = 1$.

But in this case, the two series-connected integrators become not enveloped by stabilizing feedbacks. Consequently, additional proportional feedback enveloping one of the integrators is necessary.

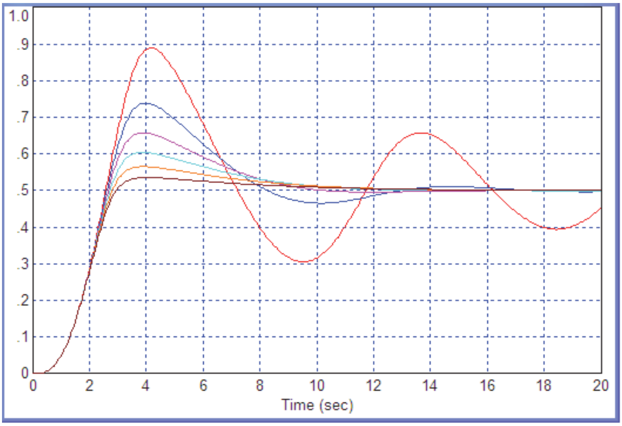


Fig. 11.14. The same values with the input shock equal to 0.5; the gain *k* is changing from the top to the bottom of the graph, respectively: 1.5; 2; 2.5; 3; 3.5 and 4

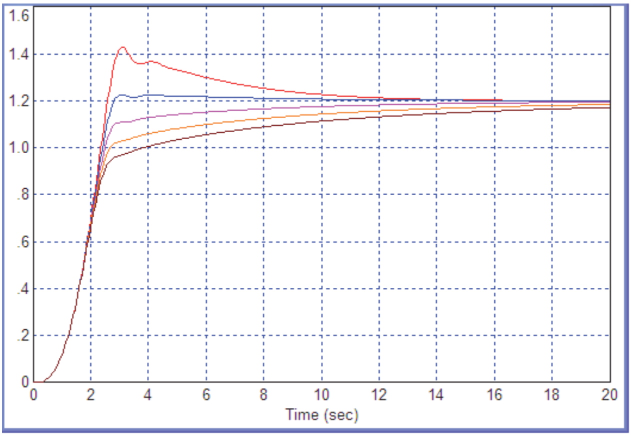


Fig. 11.15. The same values with jump at input equal to 1.2; the gain *k* is, from the top to the bottom of the graph, respectively: 1.4; 1.6; 1.8; 2 and 2.2

Fig. 11.16 shows the structure and Fig. 11.17 shows the simulation result in program *VisSim* accordingly with this variant of the proposed method. If the jump of the input signal does not exceed 1.8 units, the system is close to linear. If this value will be increased, the system becomes unstable. Namely, when the output signal value is 1.9, then rapidly growing amplitude oscillations occur (see. upper graph in Fig. 11.17).

Fig. 11.18 and 11.19 show the results of attempts to find better values of the coefficient of proportional feedback (parallel to compensatory nonlinear loop). Both increase and decrease of this ratio does not increase the stability of the system.

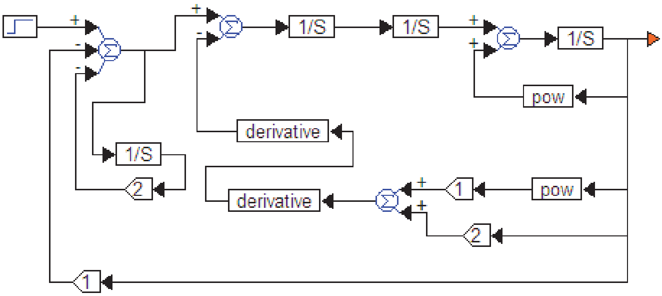


Fig. 11.16. The structure for the simulation of systems with almost full compensation of nonlinearities

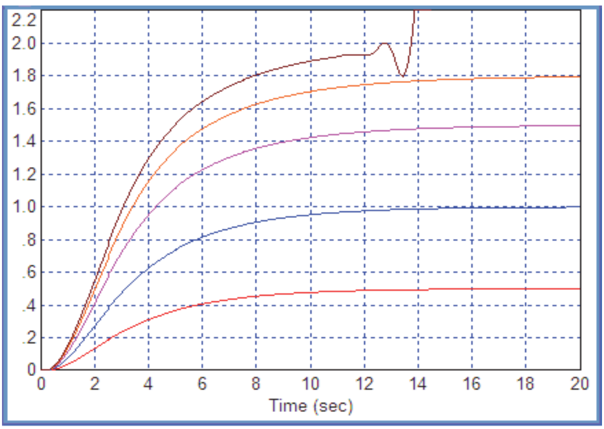


Fig. 11.17. The result of the simulation of systems according Fig.11.16: transient processes for small values of the input signal are corresponding to jump in a linear system with high quality control; problems arise only when the input jump is 1.8 or more

Fig. 11.20 and 11.21 show the result of reducing the integration step from 0.1 *s* to 0.01 *s*. It can be seen that the stability of the system has been raised. Now the system has become stable with the jump of 2.0 units, as well as with any smaller value of this jump. The reason of it this restoration of stability is better fidelity of the model of compensating loop to the model of the initial compensated loop. The compensating loop comprises two additional integrator and two additional differentiating units. Each integrator introduces a delay which value is the integration step.

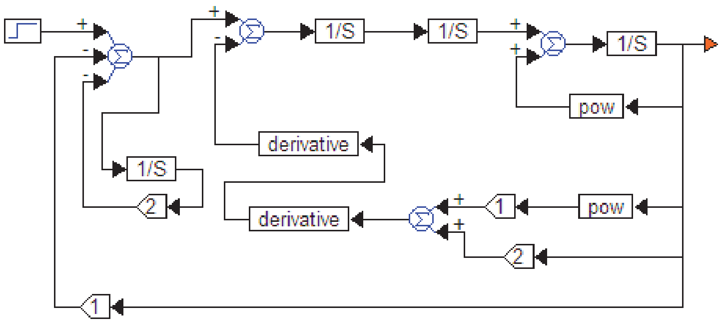


Fig. 11.18. Structure with attempts to find other feedback coefficients which do not lead to success (see transient processes in Fig. 11.19)

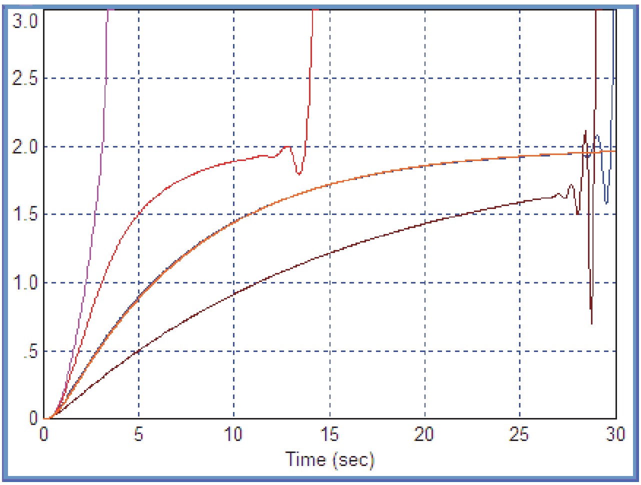


Fig. 11.19. Transient processes to Fig. 11.18

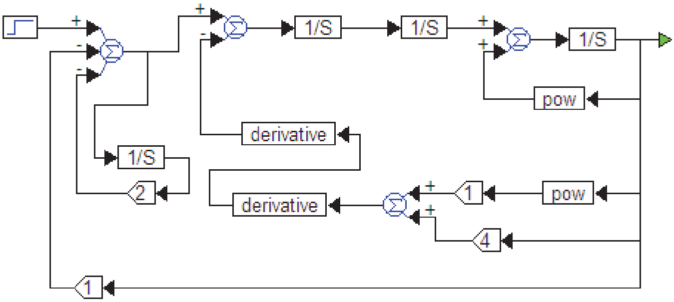


Fig. 11.20. Modified coefficients in structure according 11.18

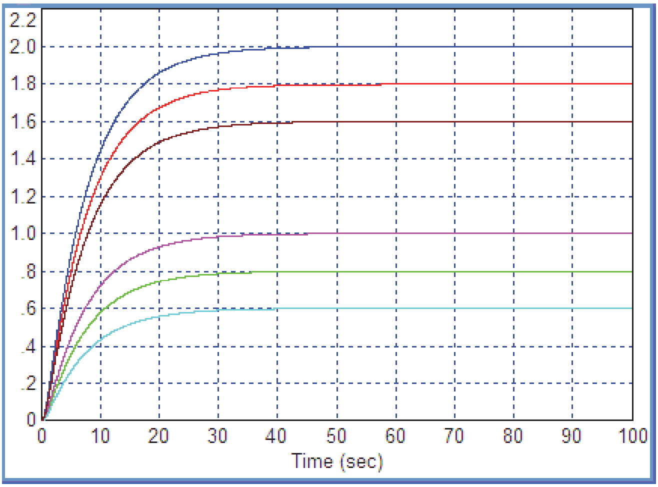


Fig. 11.21. Transient processes for structure of Fig. 11.10: illustration that the success occurs at lower values of the integration step from 0.1 s to 0.01 s, and the linearity of the system is achieved when the input jump is 2 units or less

Thus, in the compensating circuit implicitly contained delay link by an amount equal to twice of the integration step.

This rule is confirmed by further modeling. With further increase in size of the jump up to 4 units, the stability of the system becomes broken again, as Fig. 11.22 demonstrates with the transients processes in it. But it is enough to reduce -integration step again from 0.01 *s* to 0.001 *s*, and the stability of the system for these values of the input signals is again restored, as shown by the graph in Fig. 11.23.

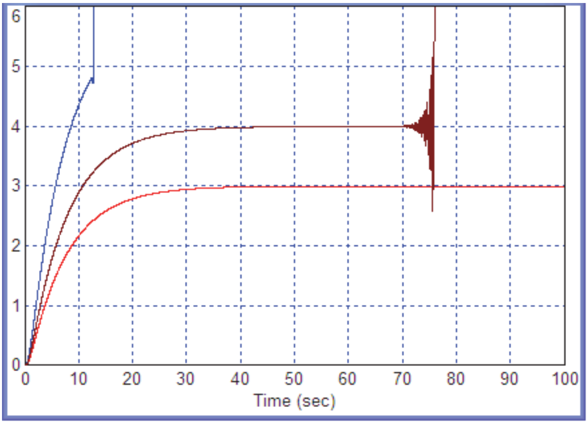


Fig. 11.22. Violation of stability with increasing input jump to a value of 4 units

Therefore, in the case of the implementation of the control system based on digital technology, ADC and DAC is required to ensure the best stability of the devices on the highest possible speed.

However, even this does not solve the problem completely for the following reasons:

1. These integrators are included into the object model, and increasing of the regulator speed does not change the speed mismatch problem of the problematic loop to compensate its influence fully.
2. The model of object can be known with insufficient accuracy.
3. The object model coefficients may vary with time during its operation.
4. The implementation of coefficients can also be carried out only within a given errors, even with quite small one.

Therefore, for all kinds of calculations it is useful endeavor to provide robust control, i.e. the calculation of such regulators, in which the system remains stable even in the case of a small deviation of the true parameters of the object from the values of these parameters, used in the calculation of the regulator.

For the considered in the last section problematic case, the design of robust regulator can be fundamentally impossible.

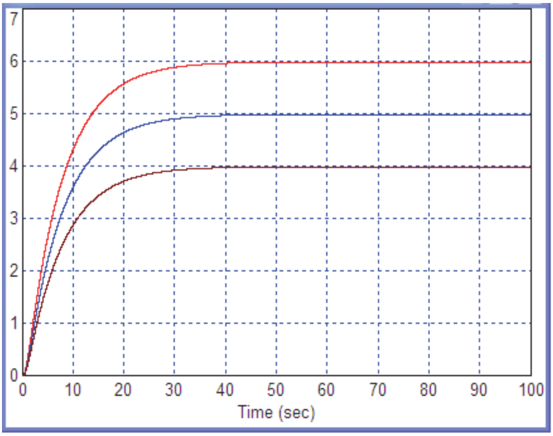


Fig. 11.23. Restoring of the stability when changing the integration step from 0.01 *s* to 0.001 *s*

The reason of this problem could be understood by considering the static state and the system.

To the situation when the output signal of the object would be in steady state, i.e. not changed (and in this case it would be equal to the input specified value), at the third (last) integrator signal must be equal to zero.

However, signal equal to the cube of the output value of the object is applied through the adder and the nonlinear element to the input of the integrator signal. Therefore, to compensate for this quantity, the output signal of the second (penultimate) integrator must also be equal to the cube output magnitude but with a minus sign. And with the aim of keeping by this integrator of its output value in the static mode, the output signal of the first integrator must also remain zero at the end of the transition process. Also at the end of the transient process zero control signal of the regulator output must take place. This is necessary but not sufficient condition. Furthermore, these signals are related by relations, according to which some of these signals are the integrals of the other or the differences of the other signals in agreement with a mathematical model of the object. The respective signals are shown in Fig. 11.24.

In this case, all of these signals are generated only from the shape of the control signal $u(t)$. From this it can be seen how complicated are the requirements for this signal, and thus for the regulator, which generates the signal from the prescribed value $v(t)$ and the output value $y(t)$.

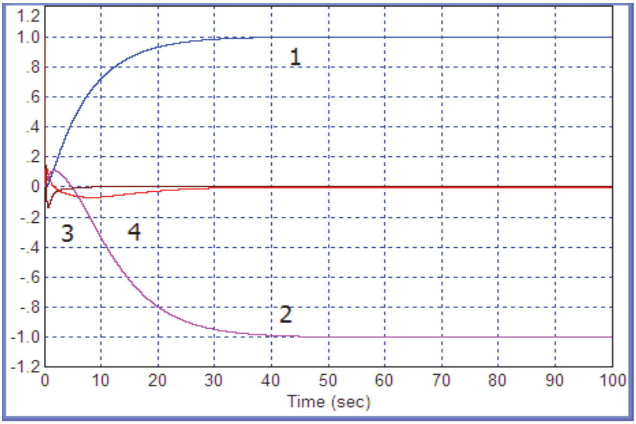


Fig. 11.24. Graphs of transient processes, including internal values of the object: 1 is output value, 2 is output of the second integrator, 3 is controlling signal, 4 is output of the first integrator

In order to investigate the robustness of the resulting the regulator gain in the compensating tract has been changed. These changes did not cause significant changes in the transient processes with the setting value equal to 12 units or less, the system remains stable. If the value is more than 13 units, it is not stable.

This is true within the frames of the limited value of the input signals (up to 12 units of reference, and correspondingly the output signal, which corresponds to 1728 units at the output of the cube link).

In this chapter, a method for design of regulator for problem objects has been proposed and examined. Object model can include many integrators and unstable internal loop. The method has been examined on the example of non-linear object with a positive feedback, which includes non-linear element elevating the input signal into a cube.

The efficiency of the systems designed by this method has been demonstrated with simulations. In case of using of the method based on the full compensation of the nonlinearity, error of the coefficient about 10% does not cause the loss of stability or a significant change in the quality of the transient process. This is true within the limited magnitude of the input signals (up to 12 reference units, and corresponding output signal 1728 at the output cubed link).

12. DESIGN OF A PIECEWISE ADAPTIVE REGULATOR

12.1. Robust system as a prototype of the adaptive system

One of the methods for calculating the robust regulator is discussed in Chapter 6.3. It consists of optimization with simultaneous modeling of a set of systems with the same regulators, but different objects, described by the most characteristic's models from the set of possible models of this object. The main idea of this approach is to ensure that identical regulators provide the stability of locked systems not only with an object with nominal parameters, but also with selected samples of models of an object with changed parameters. If in the numerical optimization of the regulator for a single object, the integral criterion of the quality of the system, is used as the criterion of optimality, then in the numerical optimization of one regulator for a set of objects, the sum of these criteria should be used. To ensure the required quality of the system for various values of object parameters, it is suggested to use a lot of objects described by the regulator model, each of which is characterized with different parameter values, as shown in Fig. 12.1.

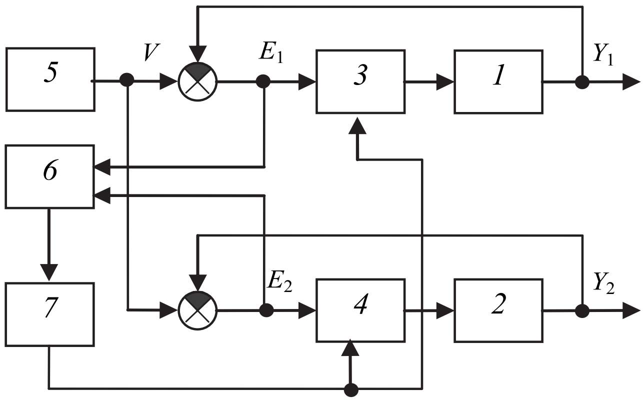


Fig. 12.1. Regulator optimization scheme: 1, 2 are models of the object with different parameter values; 3, 4 are identical models of regulators; 5 is the former of the test impact; 6 is system cost analyzer; 7 is regulator parameter optimizer

For each of these objects, an identical regulator is simulated. Its parameters are calculated by the method of optimization by the criterion, which includes the sum of the errors of all systems in the set.

Earlier in Chapter 6.3, the method of designing a robust system for an object with distinct non-stationary properties was analyzed. Features of the robust regulator are that it must provide stable control with acceptable quality, provided that the parameters of the object model change, or are not known accurately. Moreover, this success of control is achieved not at the expense of changes in the regulator model, but at the expense of finding such a universal model that would ensure the solution of the task at any possible combinations of parameters of the model of the object. Obviously, the solution to the task of successful control can be unachievable by this method.

Unlike robust regulators, adaptive regulators can change the parameters of their mathematical model depending on the current parameters of the object model. The class of problems that can be solved in this way is much wider, and the results can be much better. However, the design of adaptive regulators is much more difficult than designing robust regulators.

The main difficulty in implementing adaptive systems is, first, in determining the current model of the object, and secondly, in calculating the best regulator for this current model.

Simplification of the method for solving this problem can be achieved by splitting variants of possible mathematical models of the object into a countable set and using the robust control method within this set. In this case, the private subtask of robust control is simplified compared to trying to provide the required properties of the system with a single robust regulator. On the other hand, detailed identification of all object parameters in this case is no longer required, since it is sufficient to provide only the recognition of the characteristic features of the object model sufficient to classify the current model as one of the previously allocated classes.

Let us suppose, for example, that a control object has a mathematical model in the form of a transfer function $W_0(s)$, for example,

$$W_0(s) = \frac{k}{(Ts+1)^n} \exp\{-s\tau\} . \tag{12.1}$$

The non-stationary properties of the object consist in the fact that in some predetermined limits all the parameters of its model that enter into this function can change, namely: k is the gain, T is the time constant, n is the order of the model, and τ is the time constant of the delay link.

The output signal of the object $Y(t)$ should coincide with $V(t)$ as precisely as possible. An unknown noise is affected by the object and the parameters of the object model are slow (that is, 100–1000 times slower than the rate of change of the object output signals) and change in an unknown way in time.

If the parameters of the transfer function (12.1) change over time, then the robust regulator remains unchanged, whereas the adaptive regulator must vary depending on these changes:

$$W_R(s) = W_R(s, n, k, T, \tau) . \tag{12.2}$$

The piecewise robust regulator in our concept is a regulator whose structure (mathematical model) depends on one parameter which is the number of the subset to which the current state of the object model is assigned. In the case of a piecewise robust regulator, all the regulator's coefficients can be taken from a pre-calculated table, that is, they can be fixed from the given subset. This is much simpler than the continuous calculation of new coefficients, based on the set of newly defined parameters of the object model.

Determining the situation of assigning a model to one of the predefined classes is also much easier than determining the entire model of an object completely. As a result, two complex procedures (identification of the model of the object and calculation of the regulator model) are replaced by two simple procedures (referring the model of the object to the selected class and selecting the regulator model depending on this class).

12.2. An example of splitting a set of object parameters into subsets

One of the primitive variants of splitting the set of object parameters into subsets is the division of the range of admissible values of each of the changing parameters.

For example, if in (12.1) n is an integer in the range from 3 to 5, this automatically gives a breakdown into three subsets of this parameter. The continuously varying parameters T, k, τ can be divided into an arbitrary number of intervals. For simplicity, we can apply the partitioning of intervals into two, if there is no valid reason for another choice. If, as a result of solving the problem, it turns out that such a partitioning is not enough, one can apply a smaller fragmentation of intervals over one or several of the selected parameters.

A more intelligent approach consists of finding the patterns of joint influence of parameters on the quality of the system and, accordingly, on the choice of the regulator. For this purpose, an analytical analysis of the influence of these parameters or the combination of the resulting regulators on the basis of the results of their numerical optimization can be used. The analytical method is difficult to develop in a general form, although in some special cases this problem is solved quite easily. The combination of the results of optimization can be quite simply formalized.

For example, for $n = 1$, $\tau = 0$ in (12.1), one can see that an increase of T in m times also displaces the high-frequency part of the amplitude-frequency characteristic of the object, as a decrease of k by m times. Therefore, the whole range of admissible values of T and k can be divided according to the values of the product $G = T \cdot k$ by the desired number of intervals, preferably equal in the logarithmic scale. For example, if G_1 is the minimum value of this product, and G_4 is the maximum value, it is supposed to be divided into three intervals, it is advisable to choose two internal boundary values of the interval G_2 and G_3 by achieving the ratio:

$$G_2 / G_1 = G_3 / G_2 = G_4 / G_3 \,. \tag{12.3}$$

A similar result could be obtained if the range of admissible values of the coefficients was divided into two, as well as the interval of permissible values of the time constants. This would give four areas. Assigning the corresponding characteristic central values k_1, k_2, T_1 and T_2 to the regions with the smallest and largest values of each of the parameters, we would obtain four different pairs of combinations of values of the regulators: $Q_1 = \{k_1, T_1\}$, $Q_2 = \{k_1, T_2\}$, $Q_3 = \{k_2, T_1\}$, $Q_4 = \{k_2, T_2\}$. Further we might find that the controls for Q_2 and Q_3 are identical.

12.3. Identifying the belonging of the object model to a given subset

Methods and devices for identifying the belonging of the current set of object parameters to a specific subset can be based on the evaluation of these parameters or on a different principle, if this is more effective. In this case, test actions can be used provided they are sufficiently small so as not to disrupt the result (accuracy and quality) of the object control.

For example, the value of the static gain *k* can be approximated by giving a small step effect to the object. The order of the object *n* can be determined by measuring the transfer function in the high-frequency region, at two or more characteristic frequencies. Determination of the value of the time constant of the delay link τ can be carried out by the correlation method, that is, by feeding a pseudo-random small signal to the input of the object and by finding the delay time of this signal corresponding to its greatest correlation with the output signal of the object.

13. OPTIMIZATION OF THE REGULATOR FOR MULTICHANNEL OBJECTS (MIMO)

13.1. The task of control of multichannel object

Multichannel objects refer to special class of objects, in which there are several inputs and the same number of output values, and each input value affects each output value, which is called cross-influencing (or cross-linking). It is required to design a regulator that provides autonomous control of each output value with the help of pre-selected input quantities.

Example 39. For example, in a chemical reactor, a reaction takes place, the rate and temperature of which depend on the amount of several products fed into it. It is necessary to separately control the temperature and volume of the product in the reactor using the rates of supply of two reagents.

If the object is linear, as a rule, its mathematical model can then be set by the matrix transfer function in the operator area, that is, in the form of the Laplace transform relations from the output signals to the input signals that generate them.

The principle of the most powerful influence. The simplest approach to solving this problem is to choose a correspondence between input actions and output values based on the principle of the strongest influence, if possible.

Example 40. If the frequency and power of radiation are to be stabilized in a semiconductor laser, taking into account that each of these parameters is influenced by the temperature of this laser and the pump current, then there are two possible ways to control: a) controlling change in power by changes in current, and control changes in the radiation frequency by the changes in the laser temperature; b) controlling change in power by changes the temperature, and control changes in the radiation frequency by the pump current. In the second case, the problem posed will be much more difficult to solve, since the current has a much greater effect on the radiation power than the temperature. This statement is far from obvious; therefore it requires an explanation of how it is possible to compare four quantities with different names. Indeed, the power increment from the current is measured in *Watts per Ampere*, the power increment from the temperature in *Watts per degree*, the frequency increment from the current in *Hertz per Ampere*, and the frequency increment from the temperature in *Hertz per degree*.

We represent the matrix of static coefficients in the following form:

$$K = W(0) = \begin{bmatrix} k_{11},(W/A) & k_{12},(W/°C) \\ k_{21},(Hz/A) & k_{22},(Hz/°C) \end{bmatrix}. \tag{13.1}$$

Let multiply the left column by 1 *A*, and the right column by 1 *degree*. After that, divide the upper line by 1 *W*, and the bottom one by 1 *Hz*. We obtain dimensionless quantities. Among these quantities, we find the largest value. If it is not on the main diagonal, we change the numbering of the second index, namely: we replace the unit with two, and replace two with unit, that is, we interchange the columns. As a result, at least the largest value will be in the main diagonal. This means that at least one of the two input actions will cause the response in the forward channel to be more significant than in the secondary channel. If the second element in the main diagonal is larger than the other two elements, the problem of design of multichannel regulator will be simpler than if this condition was not met.

When this condition is met, a developer can attempt to control each value in a scalar (single-channel) control loops. For example, he can make two scalar feedbacks with the PID regulators in the forward link. In this case, the matrix regulator can be represented by a diagonal matrix in which the non-zero terms are located only in the main diagonal. This approach can be effective if the object model has the strongest influence between the selected control signals and the output values associated with them. That is, in the whole range of operating frequencies, the matrix transfer function is characterized by the fact that in the main diagonal there are terms whose magnitude is much higher than the magnitude of the terms in the rest positions of the matrix.

It may be that this approach is not effective, since the cross-channel influence is not so small that it can be neglected in the design of the regulator, that is, it is sufficiently effectively suppressed by the action of the main control loops (i.e. main diagonal elements of the matrix). In this case it is necessary to use the method of designing a matrix regulator.

In the case when the object is linear, one of the existing analytical methods for designing of regulator can be used to control such an object. For example, if the elements of the regulator transfer function are aperiodic links, then it is possible to offer a matrix that is an inverse of the regulator matrix multiplied by the coefficient and integrator, that is, the factor *K*/*s* as a regulator. In this case, the result of multiplying the matrix of the regulator by the matrix of the object is a diagonal matrix with integrators in the main diagonal. Such a system will successfully solve the problem of autonomous control of output values.

If the elements of the object are of higher order, a developer can try to approximate them with first-order elements and apply this method, but the result may not be so successful. There are other methods of analytical calculation of multi-channel regulators. However, if the object contains delay links and (or) non-linear links, then there are no analytical methods for calculating the regulators.

Numerical methods use simulation of the system including the object and regulator models. Based on the analyzing of the simulation results for various numerical parameters of the regulator, software selects parameters for which the problem is most successfully solved. Numerical methods can be divided into methods of empirical selection, optimization methods and others.

Empirical selection implements some simple rules or algorithms for finding such regulator parameters that provide a satisfactory solution of the problem, even if this result is not optimal.

Optimization methods require the formulation of the quality criterion of the system in numerical form, for example, as a cost function. Then, procedures are used to find such regulator parameters that ensure the least value of this cost function.

Other methods include, for example, the Monte Carlo method, which consists in a series of experiments with the random setting of the regulator parameters, with the choice of the result that most closely matches the control objective.

Also, other methods can include all other methods based on the experience or intuition of the developer, or in common sense.

Example 41. Let us consider the task of controlling the water level in the pool and the water temperature there, provided that hot and cold water is injected into it and the uncontrolled flow rate of water and uncontrolled cooling. It is obvious that the contribution of the increase in the rate of inflow of hot and cold water will be commensurately affected both by the temperature change and by the change in the water level. There is competition between the two control loops. However, common sense suggests that to control the temperature, instead of choosing one of the two actions, increment of the flow rate of hot water or increment of the flow rate of cold water, a developer can choose the increment of the sum of these velocities, and to control the temperature, select the increment of the difference of these velocities. Such an adjusted object will be controlled more efficiently.

Indeed, let, for example, all transfer functions be commensurable. Then the increment of the level Y_1 is determined by the relation $Y_1 = W_{11}U_1 + W_{12}U_2$, the increment of the temperature Y_2 is determined by the relation $Y_2 = W_{21}U_1 - W_{22}U_2$. Let denote the output signals of the controller, respectively, R_1 and R_2. we can set the following relationship: $U_1 = R_1 + R_2$, $U_2 = R_1 - R_2$. Hence we obtain:

$$Y_1 = W_{11}(R_1 + R_2) + W_{12}\,(R_1 - R_2),$$

$$Y_2 = W_{21}(R_1 + R_2) - W_{22}\,(R_1 - R_2),$$

or

$$Y_1 = (W_{11} + W_{12})\,R_1 + (W_{11} - W_{12}\,)R_2,$$

$$Y_2 = (W_{21} - W_{22})\,R_1 + (W_{21} + W_{22}\,)R_2.$$

For example, in the case of the equality of all transfer functions, we obtain a complete diagonal decoupling:

$$Y_1 = 2W\,R_1, \;\; Y_2 = 2\,W\,R_2.$$

The considered example is easily understood on the intuitive level, but the applied algorithm of actions is difficult to formalize, therefore in the future we will assume that all possible procedures for simplifying the control problem have already been applied, but the problem is still not simplified so that it would be possible to neglect the indirect influence of the contours. It remains a multi-channel problem of regulator synthesis.

13.2. The statement of the problem and the solvability conditions

Let the object be described by a matrix transfer function of dimension $N \times N$, the elements of which are serially connected aperiodic links and links of pure delay.

It is necessary to develop a procedure for optimizing the regulator, which would allow to calculate the regulator in the form of a functional matrix, also of dimension $N \times N$, whose elements would be physically realizable links in the form of rational fractions, for example, PID-regulators.

The problem is solvable if the matrix of the object is not degenerate. If this matrix is degenerate, that is, its determinant is zero, then the problem can not be solved.

Regarding the functional matrix, the concept of degeneracy can formally not be fulfilled, but the determinant may turn out to be small, or zero in the region, for example, only of low frequencies (or zero frequencies). This means that the problem does not have a static solution, although in dynamics it can be solved.

Example 40. Let consider an example of an unsolvable control problem for a multichannel object. Let the transfer function of the object has the following form:

$$W(s) = \begin{bmatrix} \frac{2}{3s+1} & \frac{1}{3s+1} \\ \frac{6}{5s+1} & \frac{3}{5s+1} \end{bmatrix}. \tag{13.2}$$

The first column of the matrix (13.2) is proportional to the second column. Therefore, the output signals of such an object are connected, it is impossible to provide autonomous control of the output values.

Example 41. Let consider an example of an unsolvable control problem for a multichannel object. Let the transfer function of the object has the following form:

$$W(s) = \begin{bmatrix} \frac{2}{2s+1} & \frac{1}{3s+1} \\ \frac{6}{5s+1} & \frac{3}{7s+1} \end{bmatrix}. \tag{13.3}$$

The first column of the matrix (13.3) is not proportional to the second column. Therefore, the output signals of such an object are not uniquely linked and it is possible to provide autonomous control of the output values. However, in the static mode (as $t \to \infty$), the matrix is transformed into a number matrix, $s = 0$, and this numerical matrix is degenerate.

$$W(0) = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}. \tag{13.4}$$

This means that the system can not stay in an arbitrary desired steady state with the steady state of the control signals. This can be achieved only by continuously changing the control signals. In some cases, such a decision may be considered satisfactory; in other cases this is unacceptable.

13.3. Methods for solving the problem

If the problem is solvable (the matrix of the object is not degenerate), the proposed procedure for its solution contains the following steps.

1. If any simplification methods are possible, similar to those discussed in Chapter 13.1, it is advisable to apply them in the preliminary stage by adding an appropriate matrix at the input of the object. Next, the resulting modified object is considered as a new object and a developer applies all subsequent procedures to it.

2. By the elements of the main diagonal of the regulator matrix, a developer chooses PID-regulators in the form: $W_{ii}(s) = K_{Pii} + K_{Dii}\, s + K_{Iii} / s$. Here K_{Pii}, K_{Dii} and K_{Iii} are the coefficients of the proportional, derivative and integrating links, s is the argument of the Laplace transform.

3. By the other elements of the regulator matrix, a developer chooses PD-regulators in the form: $W_{ij}(s) = K_{Pij} + K_{Dij}\, s$. If the transient process is not satisfactory, the integrating terms K_{Iij} / s should also be introduced into the rest elements too.

4. A developer sets the duration of the simulation to NT, where T is the expected duration of the transient process in the system (this value can be refined during the simulation).

5. At the first input in the simulation, a developer forms the input action V_1 in the form of a step jump unit function without delay. On the second input, he forms a similar action V_2 with a delay equal to T, and so on, for the i-th input delay is $(i–1) \times T$.

6. A developer should form the weight function F as a series of sawtooth signals linearly increasing from zero to an arbitrarily given constant with a period equal to T.

7. Using N subtractors, a developer forms a column of control errors E_i equal to the difference between the prescribed signals V_i and the output values Y_i. That is, $E_i = V_i - Y_i$.

8. The received errors E_i are used as input signals for the matrix transfer function of the multichannel regulator, and also for calculating the cost function. In this case, the cost function is calculated by integrating the sum of the modules of all errors multiplied by the weight function F.

9. A developer should use the optimization procedure of the regulator and analyze the result.

10. If the result is not satisfactory, a developer can use *Smith* multichannel predictor [23] to form a new “composite object”. The above procedure to items 2–9 can be applied to the received “composite object”.

11. Instead of the *Smith* predictor a developer can use its prolongation in the form of a bypass channel 23]. The important difference between the bypass channel and *Smith* predictor is that its structure can vary.

The multi-channel *Smith* predictor is formed according to the same principle by which the one-channel *Smith* predictor is supposed to be formed, extending this principle to a multi-channel object.

13.4. The efficiency of the completeness of a PID-regulator when controlling a multi-channel object

In a series of papers on numerical optimization of regulators [21–42], techniques and methods that allow obtaining numerical parameters of regulators from a known model of the object are considered. In this case, when using a multichannel object, it is recommended to use an incomplete PID regulator, namely: the integrator is recommended to be used only in the forward control loop, that is, in the main diagonal of the matrix transfer function of the regulator. In other elements, the use of PD regulators is recommended.

This approach is justified from the standpoint of the following considerations:

1. The presence of an integrator in the main diagonal is theoretically sufficient to provide astatic control.

2. This simplification will make it possible to exclude several unknown coefficients. For example, if a developer controls an object of 2 × 2 dimension, instead of twelve coefficients, it is sufficient to find only ten. When controlling an object of 3 × 3 dimension, instead of 27 coefficients, it is sufficient to find only 21 coefficients.

In this chapter, we investigate the feasibility of such a simplification, for example, an object with close parameters of the elements of the matrix transfer function. Such an object model is one of the most problematic options, especially if there is a delay.

Example 42. Let the transfer function of an object be given by the product of two matrices:

$$W_1(s) = \begin{bmatrix} \dfrac{2}{s+1} & \dfrac{1}{2s+1} \\ \dfrac{1,5}{3s+1} & \dfrac{2}{s+1} \end{bmatrix}, \qquad (13.5)$$

$$W_2(s) = \begin{bmatrix} \exp(-0,5s) & 0 \\ 0 & \dfrac{1}{0,2s+1} \end{bmatrix}. \tag{13.6}$$

Figure 13.1 shows the block diagram of such an element when modeling it in *VisSim*.

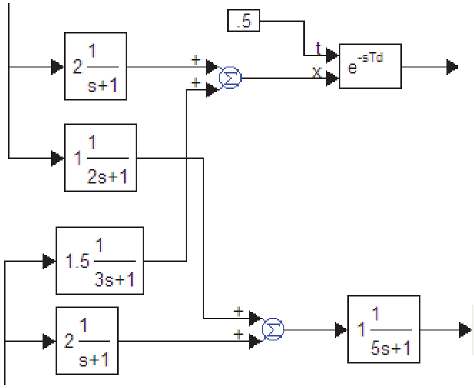


Fig. 13.1. Object structure

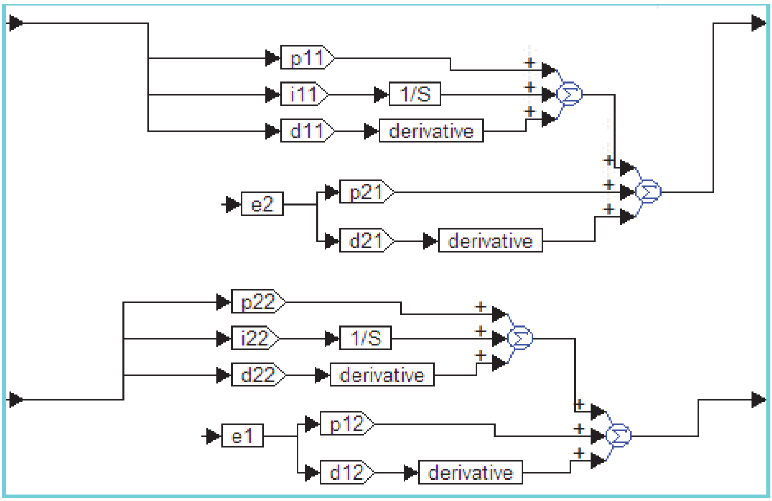


Fig. 13.2. The structure of the regulator (PID in the main diagonals and PI in the rest)

Using numerical optimization methods in software *VisSim*, we will investigate the control possibilities when using two variants of regulators, namely: regulator with an incomplete structure, shown in Fig. 13.2, and regulator with the complete structure shown in Fig. 13.3.

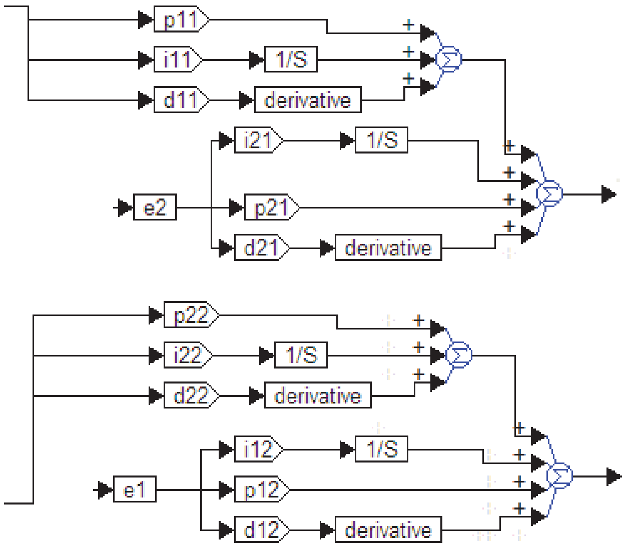


Fig. 13.3. Structure of the complete PID regulator

The example with a simplified structure was studied in detail in [42]. The result is shown in Fig. 13.4, which shows the structure for modeling the entire system.

Figures 13.5 and 13.6 show the transient processes with a separate supply of stepped input actions with a time difference equal to half the simulation time, that is, 20 *s*.

The overshooting in the first channel is from 10 to 35%. The overshooting in the second channel is about 10% when the step input is applied to the first input, that is, in this case we are talking about the cross effect of the first channel on the second one.

With other regulator parameters, a smaller overshoot value in the channels can be provided, but the required static accuracy is not ensured, in particular, in the second channel, in this case the static error is too large and increases with time.

Figure 13.7 shows the results of optimization of the same system using the complete structure of the PID regulator.

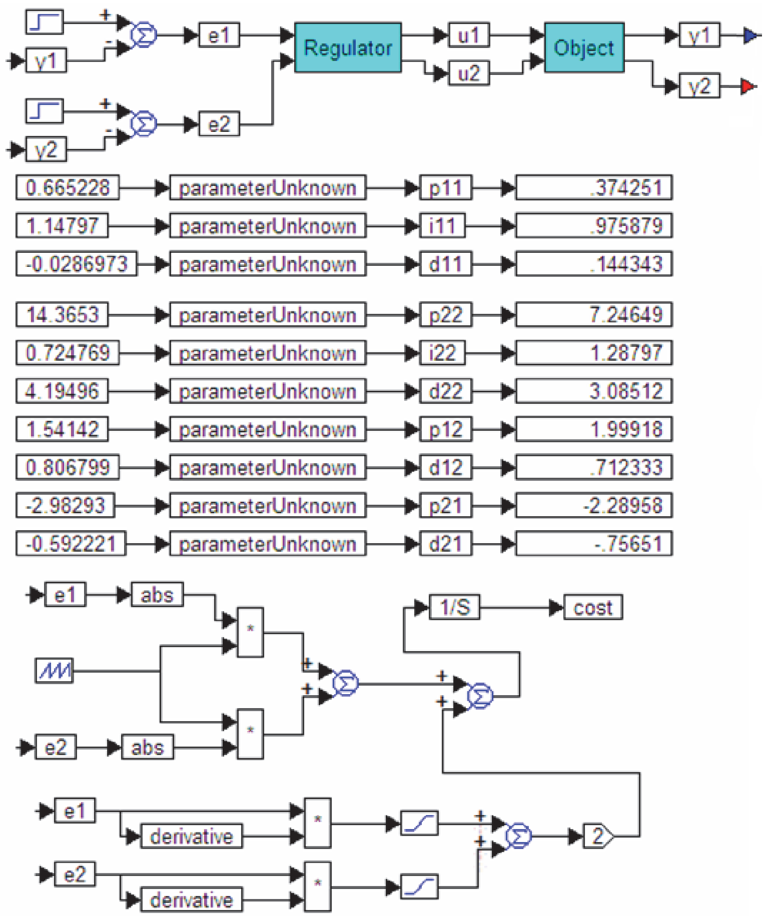


Fig. 13.4. The best result obtained with a simplified structure of the PID regulator [43]

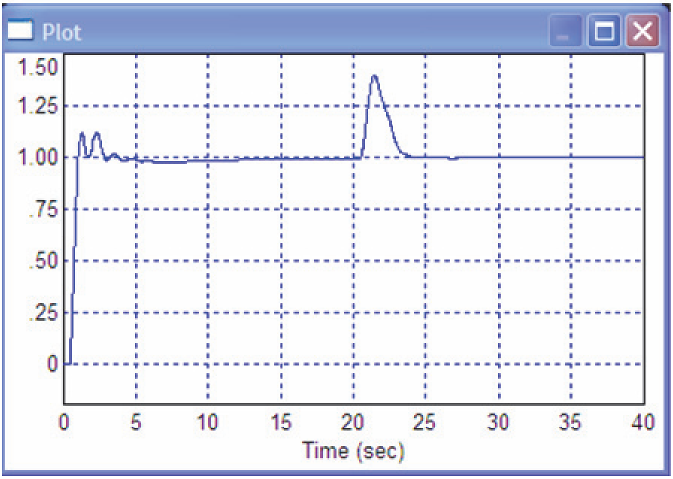


Fig. 13.5. The best result obtained with a simplified structure of the PID regulator [43]: the output signal of the first channel

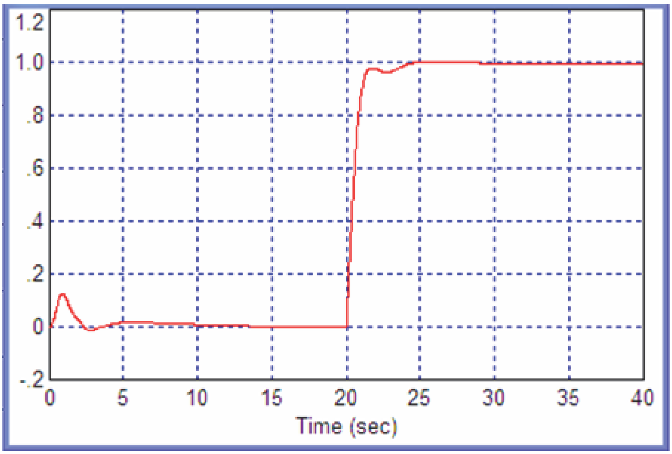


Fig. 13.6. The best result obtained with a simplified structure of the PID regulator [43]: the output signal of the second channel

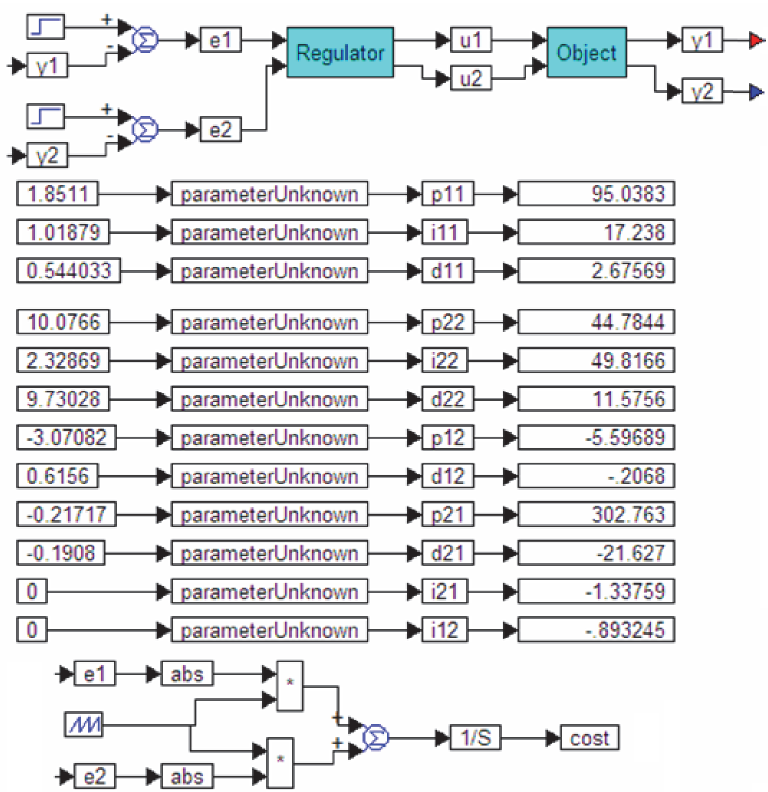


Fig. 13.7. Results of system optimization when using a complete PID regulator

Figures 13.8 and 13.9 show the transient processes at the outputs of the object shown in Fig. 13.4–13.6. As can be seen from Fig. 13.8 and 13.9, due to optimization, high speed is provided, but large overshooting, namely: in the first channel it exceeds 150%, in the second channel about 80%.

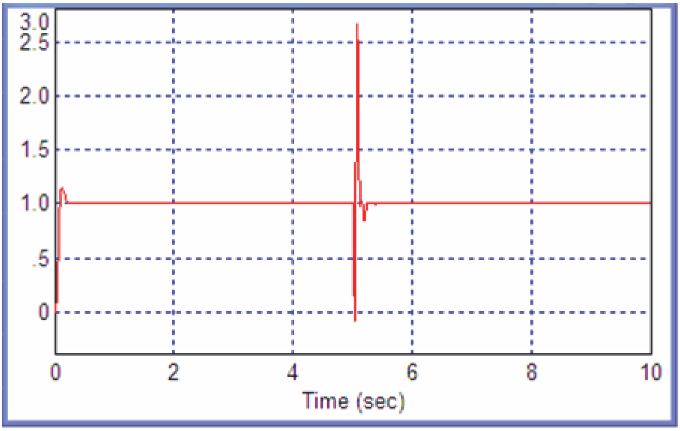


Fig. 13.8. The results of system optimization when using a complete PID regulator are the output of the first channel

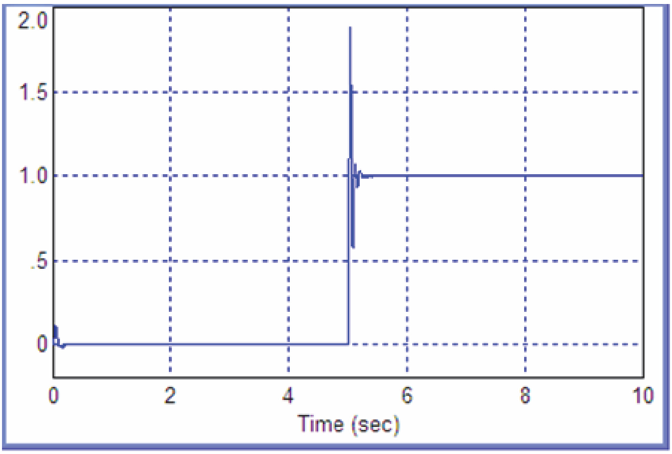


Fig. 13.9. The results of system optimization when using a complete PID regulator are the output of the first channel

This result, shown in Fig. 13.7–13.9, was obtained without using the correctness detector [23], in contrast to the result shown in Fig. 13.4–13.6. Therefore, in order to correctly compare the two types of regulators, in the second case it is also necessary to apply the correctness of motion detector, that is, the term in the objective function, which increases sharply if the product of any of the errors by its derivative is positive. The results are shown in Fig. 12.10–12.12.

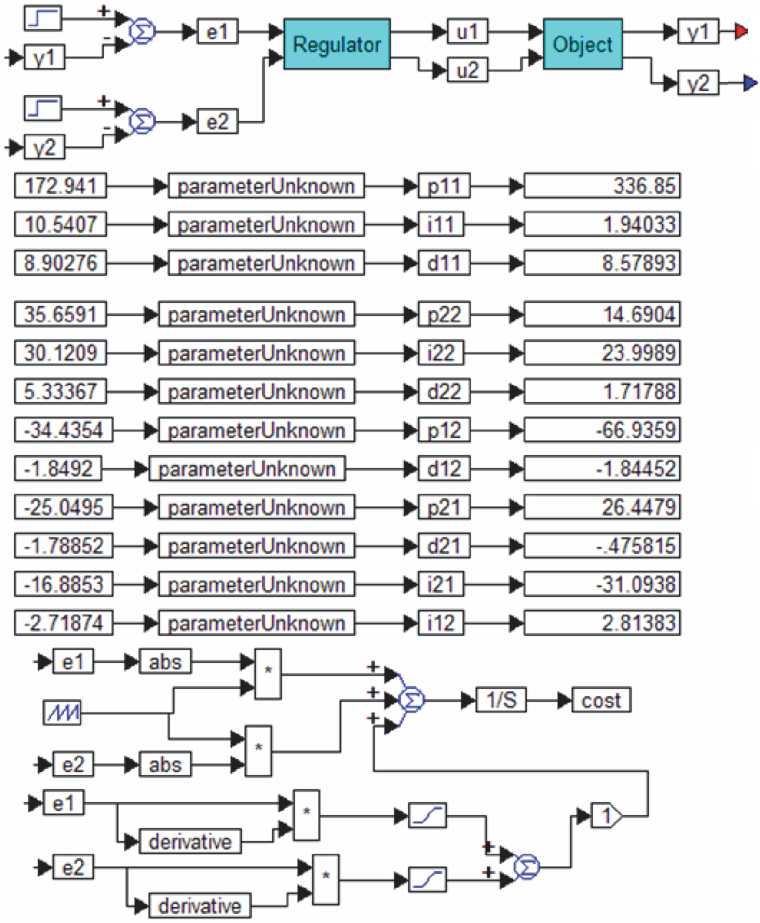


Fig. 13.10. The results of system optimization with the use of a complete PID regulator when the detector movement's correctness has been used

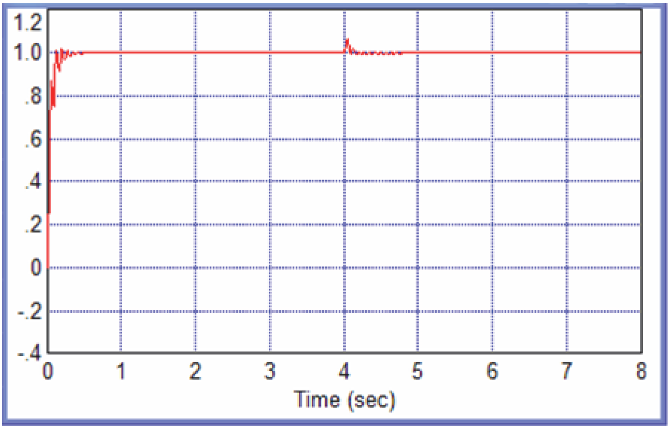


Fig. 13.11. The results of system optimization when using complete PID regulator: the first channel output

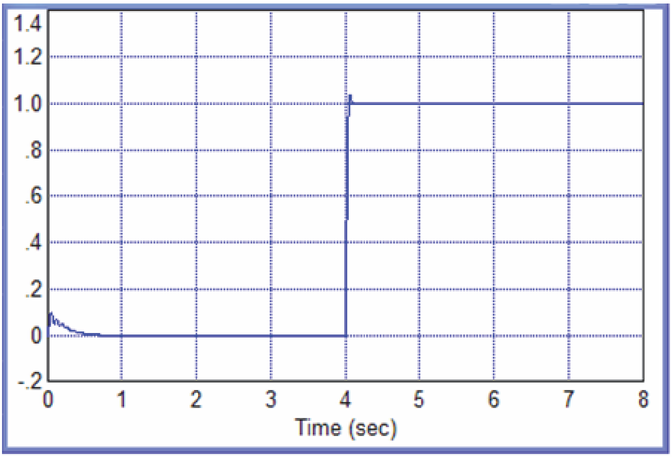


Fig. 13.12. The results of system optimization when using complete PID regulator: the second channel output

As can be seen from Fig. 13.10–13.12, reduction of overshooting is ensured to a value of the order of 2%.

Thus, it has been shown that, at least in some cases, the use of a complete PID regulator structure is more efficient than using a reduced structure.

13.5. The idea of Smith multichannel predictor

In the one-channel version, the idea of *Smith* predictor is as follows [23]. Let the model of the object consist of two parts: the minimum-phase part with the transfer function $W_{MF}(s)$ and the delay element with the transfer function $W_3(s) = \exp(-\tau s)$. In this case, *Smith* predictor is included in parallel with the object, the transfer function of which is equal to:

$$W_{YC}(s) = W_{M\Phi}(s)[1 - W_3(s)]. \tag{13.7}$$

This section explores the possibility of using this idea for the case of multi-channel object. The idea is used as a basis, which is supplemented with the above optimization technique of the multi-channel regulator.

Example 43. Let consider solution of the problem without *Smith* predictor. Let the minimum-phase part is given as a matrix with aperiodic links, that is, the elements of the matrix are first-order filters:

$$W_{M\Phi}(s) = \begin{bmatrix} \dfrac{2}{s+1} & \dfrac{0.5}{5s+1} \\ \dfrac{0.5}{5s+1} & \dfrac{2}{s+1} \end{bmatrix}. \tag{13.8}$$

We can consider an example that does not fully correspond to the approach proposed in the idea of the *Smith* predictor. Namely, let the delay element not completely be a link of pure delay, but instead there is only a pure delay in one channel, and in the other channel, there is an element with a limited speed in the form of an additional filter. In this case, such a transfer function is described by a matrix of the following form:

$$W_3(s) = \begin{bmatrix} \exp(-0,5s) & 0 \\ 0 & \dfrac{1}{5s+1} \end{bmatrix}. \tag{13.9}$$

Figure 13.13 shows the block diagram of such an element when modeling it in *VisSim*.

In accordance with the principles outlined in Chapter 2, PID regulator of dimension 2 × 2 has the form whose block diagram (also for *VisSim*) is shown in Fig. 13.14.

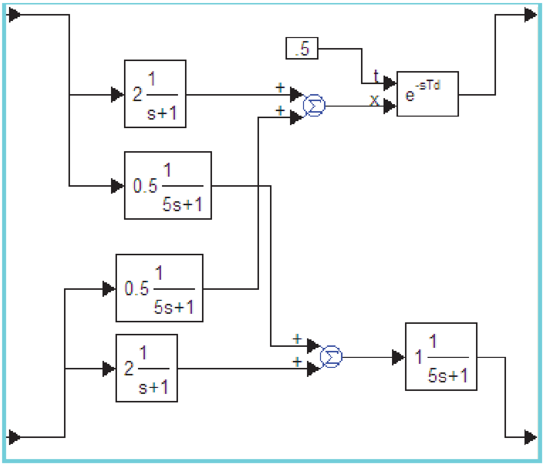


Fig. 13.13. Object structure

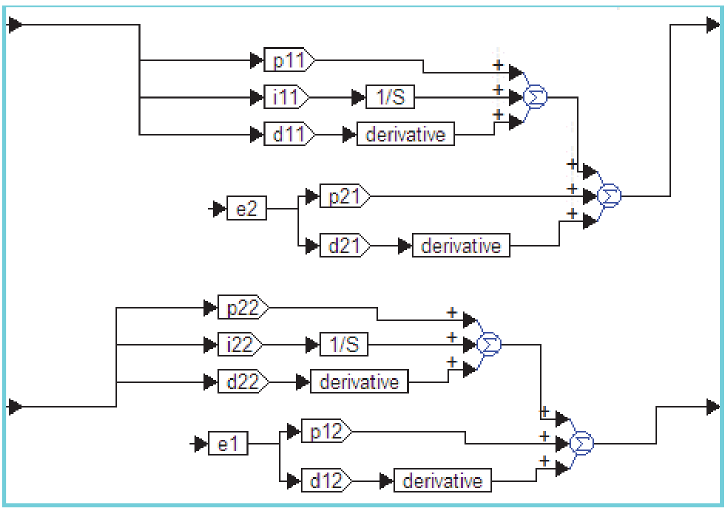


Fig. 13.14. The structure of the regulator (PID in the main diagonals and PI in the non-main diagonals)

Figure 13.15 shows the structure for modeling the entire system. On this structure, there are also blocks for optimizing the regulator coefficients, showing the results of optimization in the form of the obtained coefficients and in the form of the obtained transient processes.

Figures 12.16 and 12.17 show the transient processes. It can be seen that the overshooting on the first channel is about 30%.

In this example, the result is not as bad as to be an argument for finding alternative methods, so we will complicate the task.

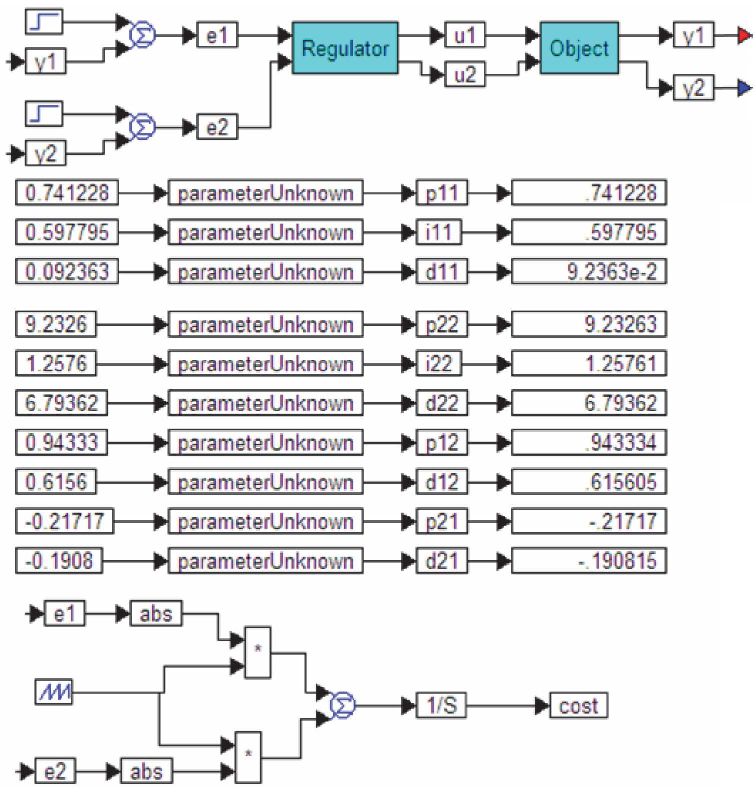


Fig. 13.15. The result of numerical optimization of the regulator: the values of the coefficients and the resulting transient processes in the system

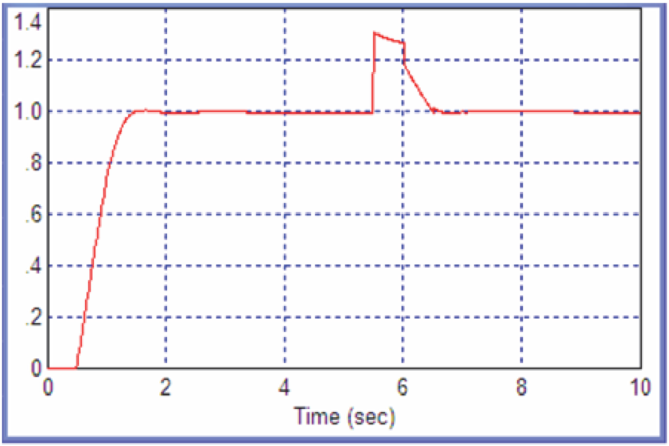


Fig. 13.16. The result of numerical optimization of the regulator: the output of the first channel

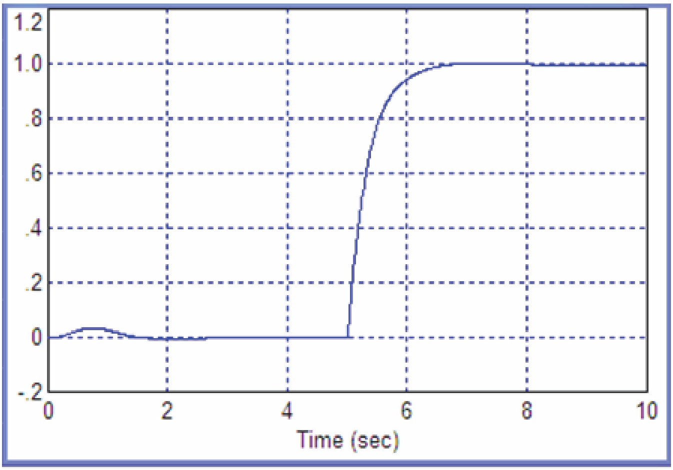


Fig. 13.17. The result of numerical optimization of the regulator: the output of the second channel

Example 43. Let consider a variant of an object of less favorable combinations of properties of the direct and side paths. Let the minimum-phase part be given by a matrix with aperiodic links, that is, the matrix elements are first-order filters:

$$W_{M\Phi}(s) = \begin{bmatrix} \frac{2}{s+1} & \frac{1}{2s+1} \\ \frac{1,5}{3s+1} & \frac{2}{s+1} \end{bmatrix}. \qquad (13.10)$$

The structure of such an object is shown in Fig. 13.18. Figures 13.19 and 13.20 show the transient processes of the corresponding transfer functions entering into the matrix transfer function (13.10) (response to a single stepwise jump action). Figure 13.21 shows the response of the object to a single jump of the first and second control inputs when they are separately fed to these inputs.

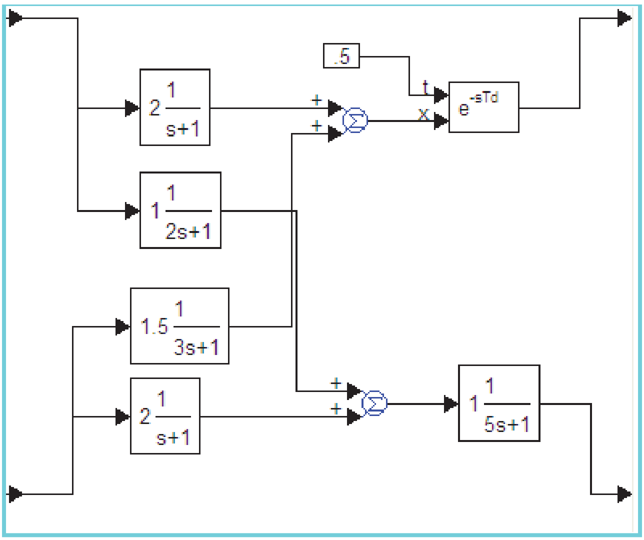


Fig. 13.18. Structure of another object: in non-main diagonals, the speed and transmission factors are raised

Figure 13.22 shows the optimization scheme and the result in the case where the first stepping action is a jump on the first channel, and then, after 20 s, a negative jump from “1” to “0” occurs on the second channel. The corresponding processes are shown in Fig. 13.23 and 13.24. It is seen that in both cases too much overshooting occurs at the output of the first channel. After the first jump, it is about 50%, and after the second – about 75%.

Also, the transient processes in Fig. 13.26 and 13.27 differ in that the time between jumps is 5 *s*. The overshooting on the second channel can rise to 125% with an unfavorable combination of input effects, which in no way can be recognized as feasible in practically any practical problem.

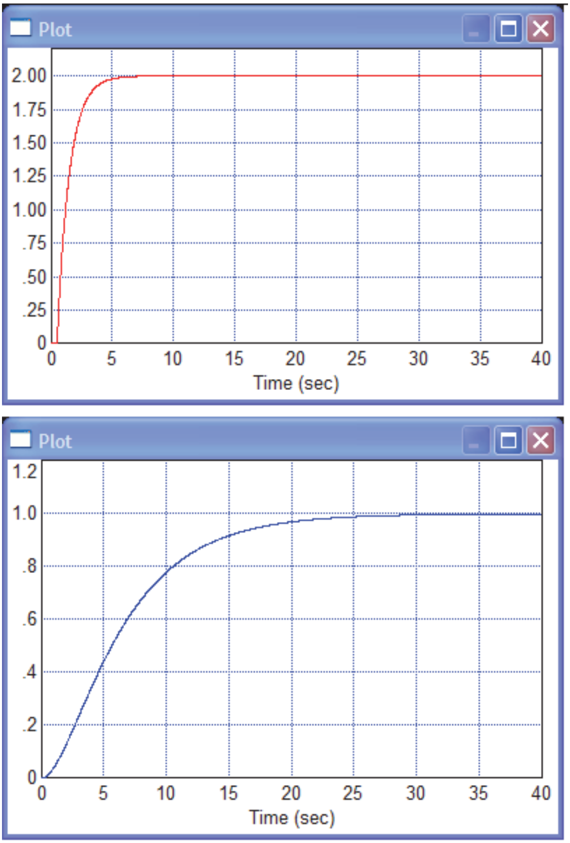


Fig. 13.19. The response of the transfer functions of the first line (13.10) to unit jump

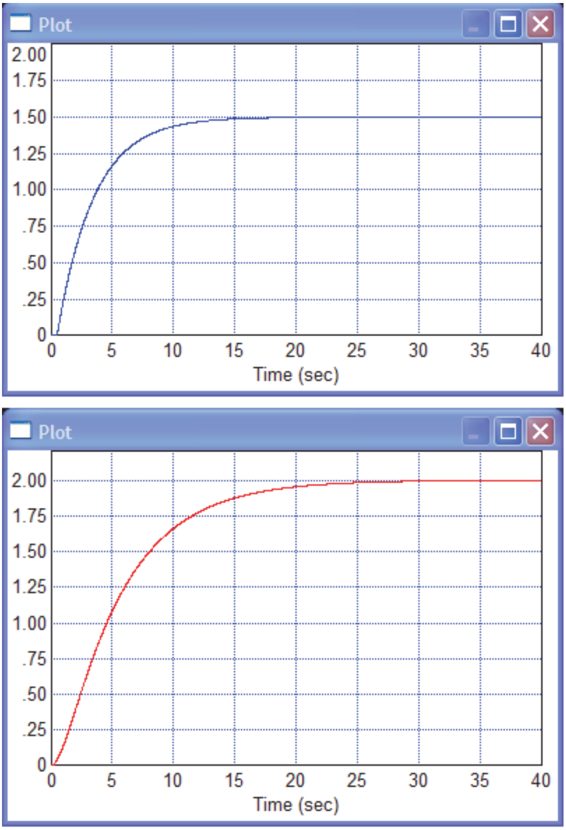


Fig. 13.20. The response of the transfer functions of the second row (13.10) to unit jump

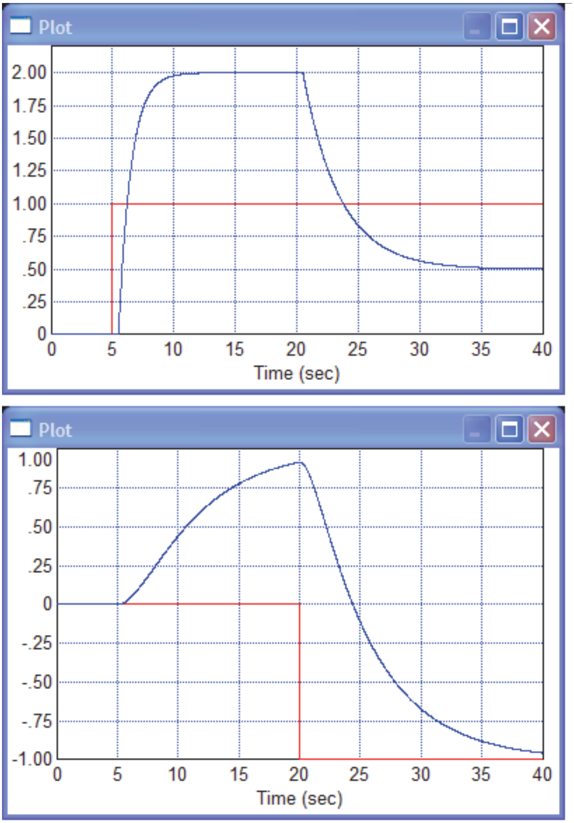


Fig. 13.21. Response of the object to unit jump of the first and second control inputs when they are separately fed to these inputs

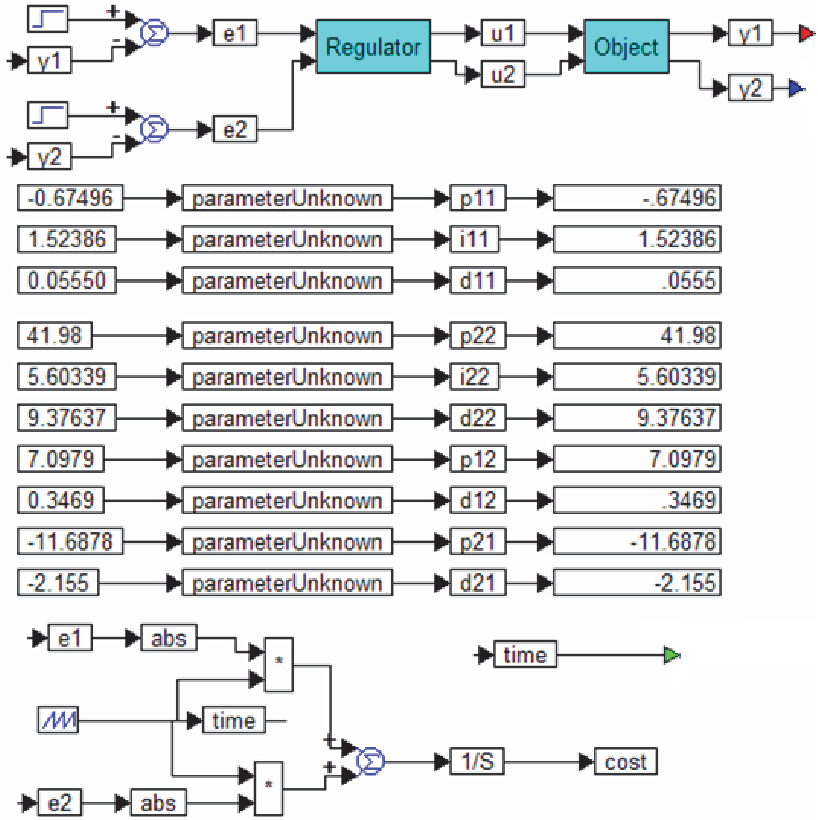


Fig. 13.22. The result of the optimization of the 2 × 2 dimension PID regulator for the object (13.10), the sequence of jumps is as following: jump at the first channel from "0" to "1", then jump at the second channel from "1" to "0", the time between the jumps is 20 *s*

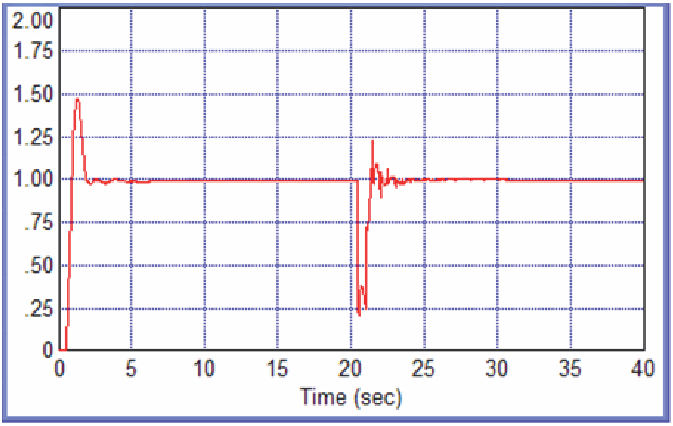


Fig. 13.23. The transient processes in the system in Fig. 13.22, the first control channel

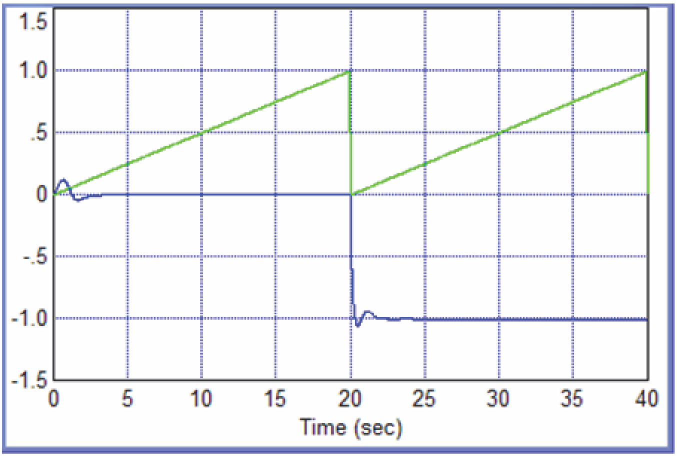


Fig. 13.24. The transient processes in the system in Fig. 13.22, the second control channel

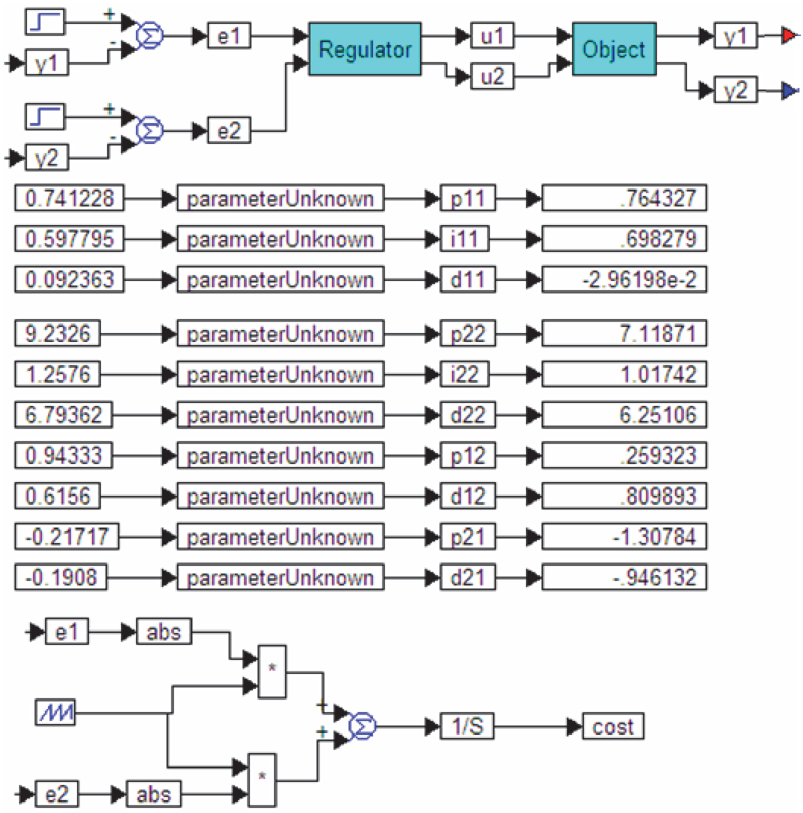


Fig. 13.25. The same result of optimization of the 2 × 2 dimension PID regulator for the object (13.10), the sequence of jumps, is the same, both jumps from “0” to “1”, the time between jumps is 5 *s*

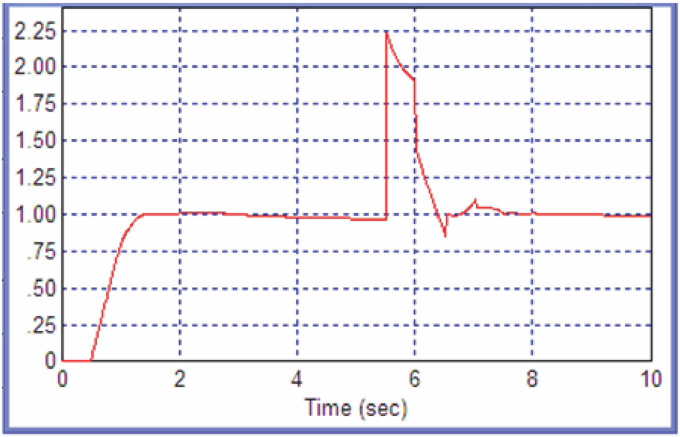


Fig. 13.26. The transient processes in the system in Fig. 13.25, the first control channel

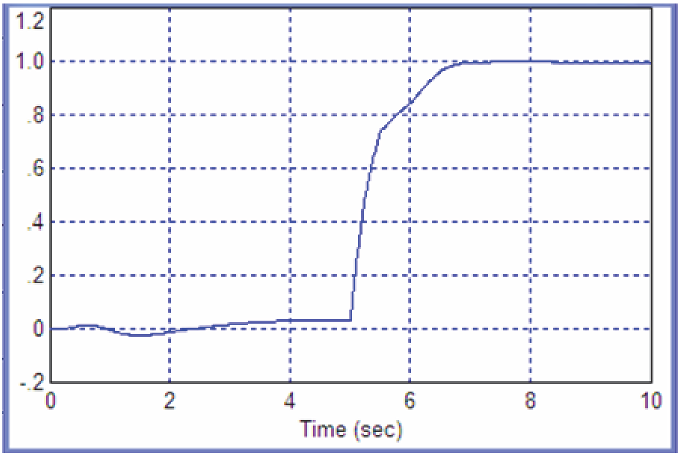


Fig. 13.27. The transient processes in the system in Fig. 13.25, the second control channel

Example 44. Figure 13.28 shows the optimization result in the same scheme for the case where a stepwise positive jump is first applied to the second channel and then, after 5 *s*, to the first channel. In the resultant system, a bad overshooting persists in the first channel.

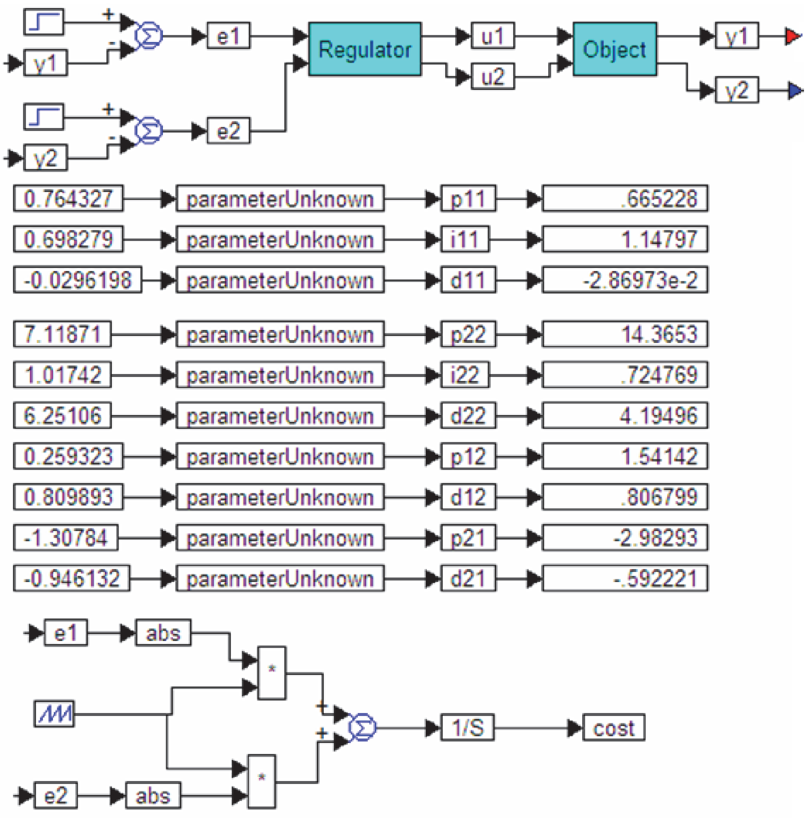


Fig. 13.28. The optimization results when first unit jump is applied to the second input, and then, after 5 *s*, unit jump is applied to the first input

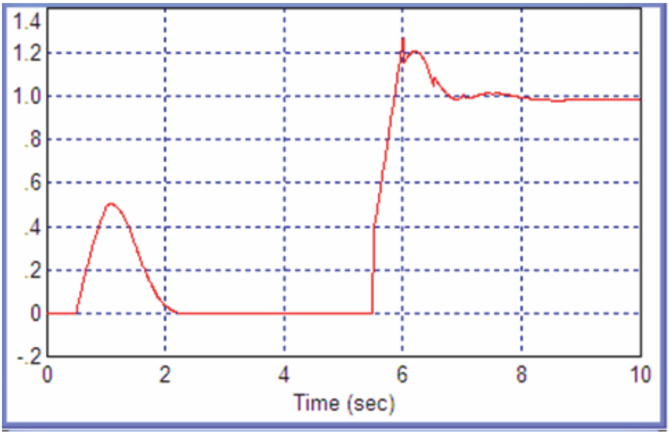


Fig. 13.29. The processes in the structure of Fig. 13.28, first channel: bad overshooting in the first channel

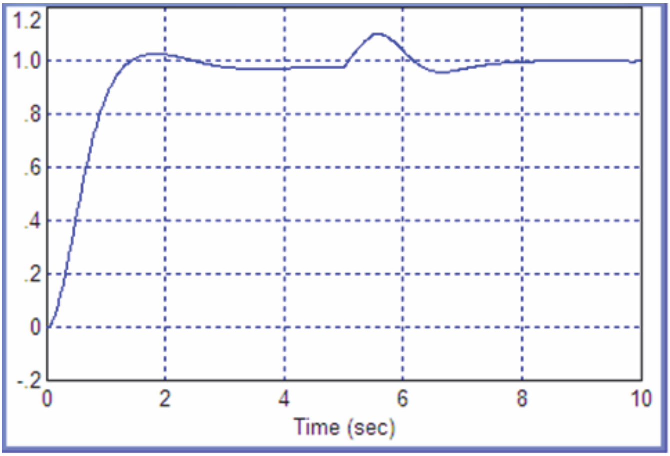


Fig. 13.30. The processes in the structure of Fig. 13.28, second channel: acceptable overshooting on the second channel

Example 45. Finally, a developer can try to increase the time between jumps and in this case; the result is shown in *Fig*. 13.31. In the resulting system, there is still an excessively large overshoot in the first channel.

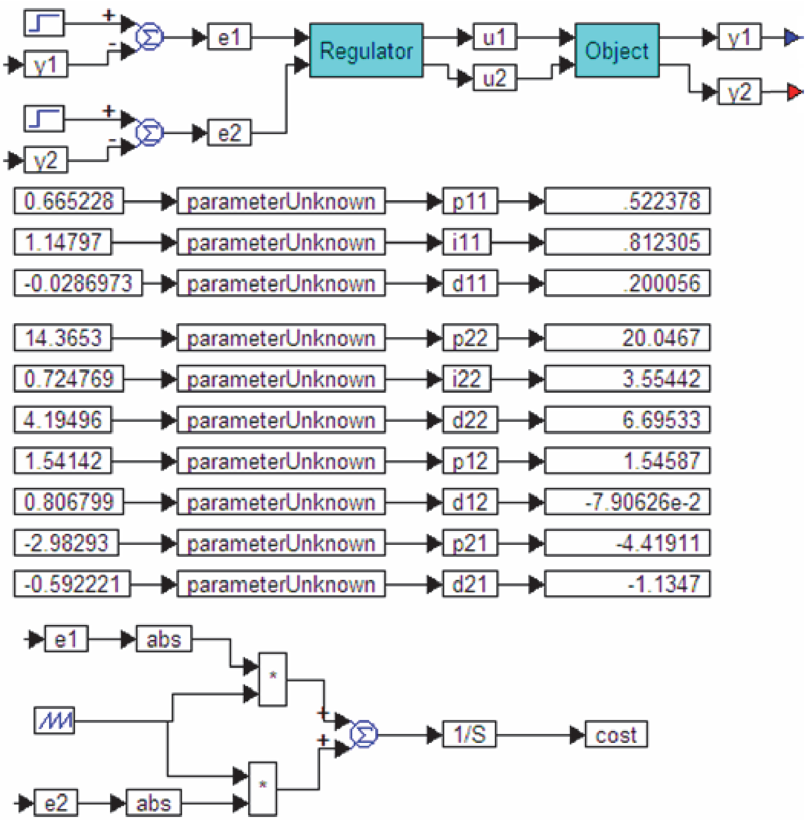


Fig. 13.31. The results of optimization when feeding unit jump to the second input first, and then, after 10 *s*, unit jump to the first input

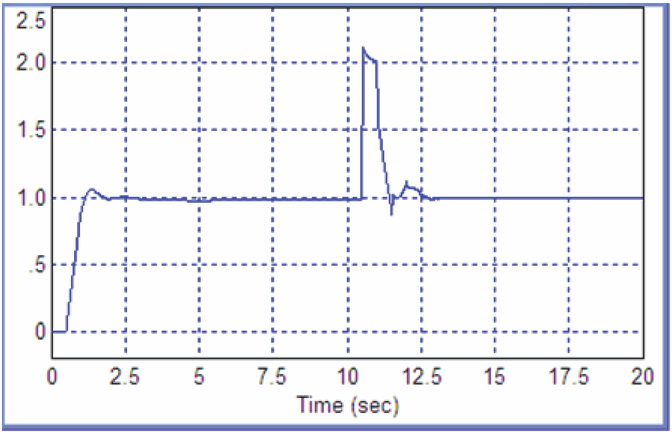


Fig. 13.32. The processes in the structure of Fig. 13.31, first channel: bad overshooting in the first channel

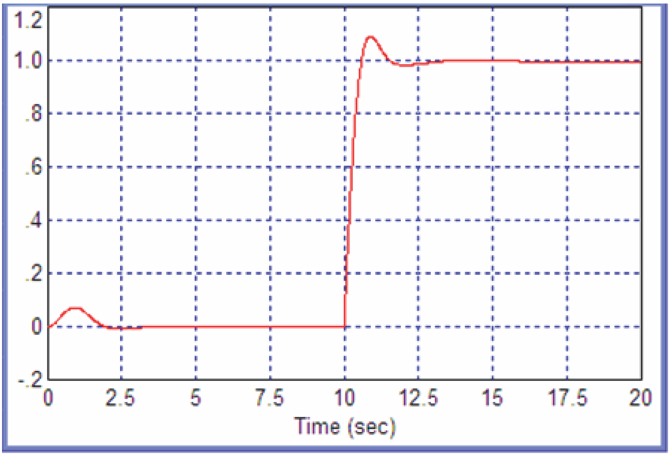


Fig. 13.33. The processes in the structure of Fig. 13.31, second channel: acceptable overshooting on the second channel

The following preliminary conclusions can be done:

Conclusion 16. The sequence and sign of the selected test actions when optimizing the multi-channel controller affect the result.

Conclusion 17. The simulation time and the time between successive test actions when optimizing the multi-channel regulator affect the result.

Conclusion 18. The multi-channel PID regulator for the object under consideration, when optimized by the criterion of the minimum of the integral of the sum of the error modules multiplied by the time since the jump arrival, did not lead to a satisfactory result: overshoot is too great.

Example 46. For more efficient optimization of the regulator, we introduce into the cost function the error growth detector discussed above. The detector determines the product of the error by its time derivative. This product with a qualitative transient process must be negative, that is, the error should decrease if it is positive, and it should increase if it is negative. In other words, the error moves to zero if this product is negative, and it moves from zero (increases in absolute value) if this product is positive.

When using a composite cost function, it is important to choose the weighting factor that determines the relationship between them. Figure 13.34 shows the results of optimization in the case when the weight coefficient for the second term is five. In this case, overshooting is effectively suppressed; however, the main purpose of control is to provide zero control error. The resulting transient processes demonstrate small control accuracy of the first channel: it is clear that the error increases with time. This result is also manifested in the fact that one of the integrator coefficients is negative (the output of block i_{22} is –1.65328). All the coefficients of the integrators of the main diagonal ($i = j$) must be positive. Ideally, only the coefficients of the side paths, that is, the coefficients whose numbering does not coincide, can be negative, for example, p_{21}, d_{21}, etc. In rare cases, one can agree with the negative coefficients of the differentiating paths in the main diagonal, but this is completely excluded for the integrator (if the coefficients of the transfer functions of the object in the main diagonal are positive).

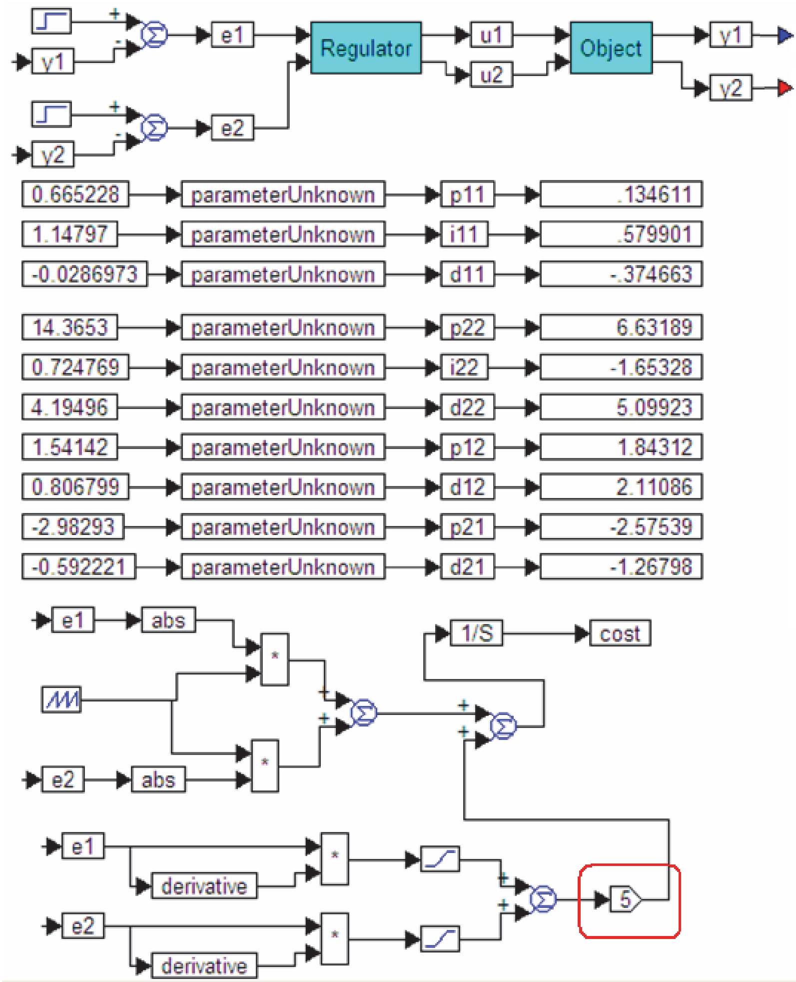


Fig. 13.34. The results of optimization when the error growth detector is introduces with a weight coefficient of 5

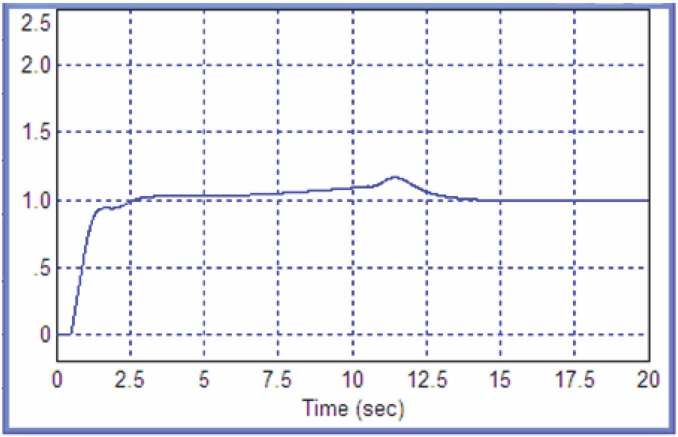


Fig. 12.35. The processes in the structure of Fig. 12.34, the first channel: a small overshoot in each channel, but poor control accuracy in the first channel

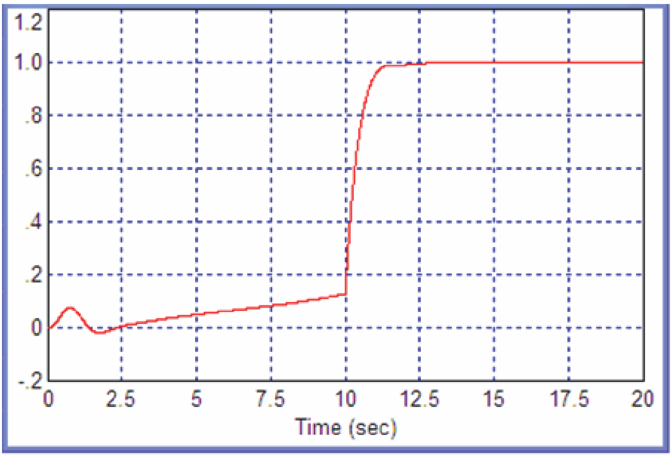


Fig. 13.36. The processes in the structure of Fig. 13.34, second channel: small overshoot in each channel, but the growth of the static error in the second channel in the first part of the process

Example 47. Figure 13.37 shows the best result obtained with this regulator structure, namely: astatism of each loop is achieved, overshooting moderately large only on one channel, astatism is confirmed by positive coefficients of integrators of the main diagonals and visually according to the graphs the assumption of astatic control is not refuted (errors tend to zero).

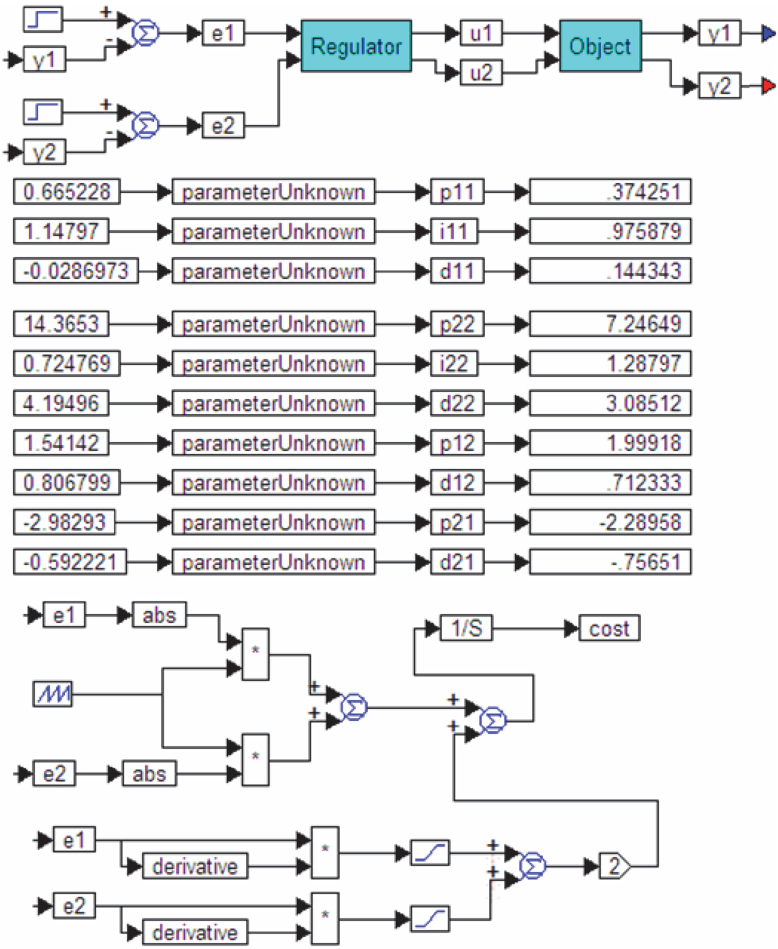


Fig. 13.37. The best result obtained with the controller structure without *Smith* predictor

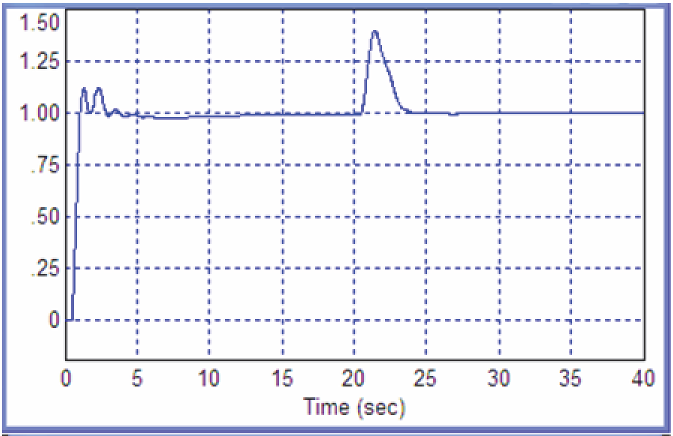


Fig. 13.38. The processes in the structure of Fig. 13.37, the first channel: overshooting is 40%

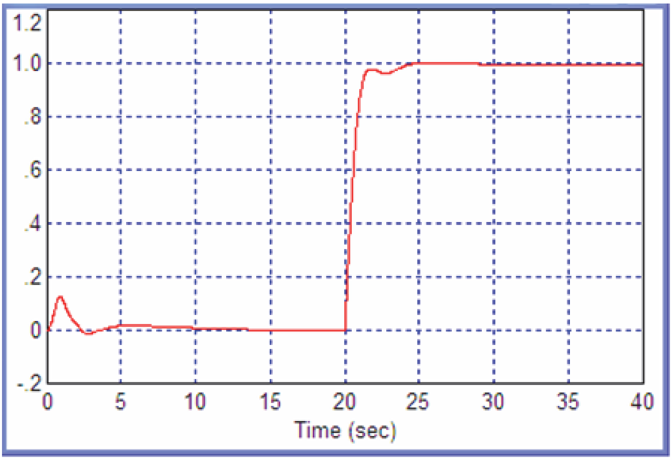


Fig. 13.39. The processes in the structure of Fig. 13.37, second channel: low overshooting is 10%

Example 48. Figure 13.40 shows the result obtained with *Smith* predictor. The structure of this predictor is also shown in Fig. 13.32. It is seen that the overshooting of the first channel is reduced to 25%, in the second channel it is somewhat less, about 20%, the static error of each channel is small enough.

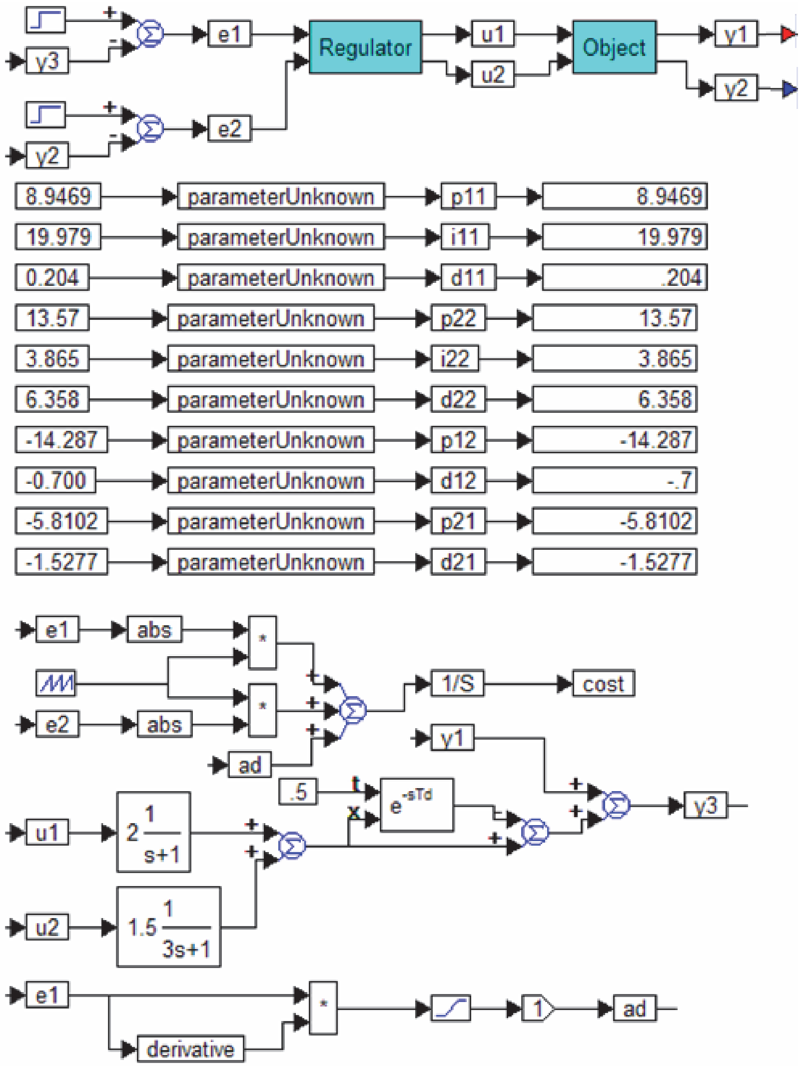


Fig. 13.40. The best result obtained with Smith predictor

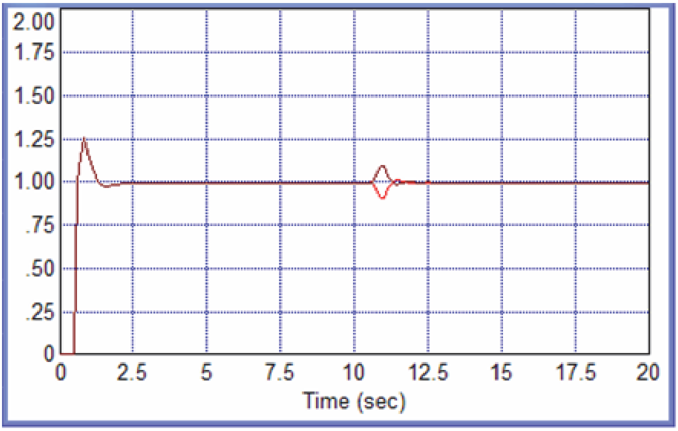


Fig. 12.41. The processes in the structure of Fig. 12.40, the first channel: overshooting is 25%

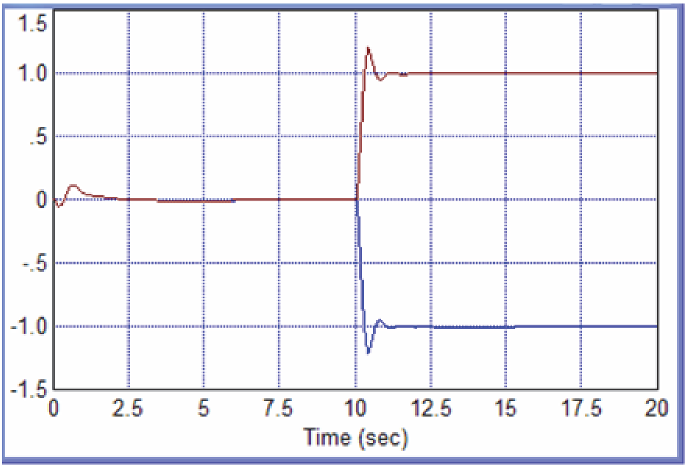


Fig. 13.42. The processes in the structure of Fig. 13.40, second channel: overshooting is 20%

Thus, the use of *Smith* predictor in combination with the proposed method of optimizing the regulator and the features of its structure gives a significant positive effect, consisting in achieving a small error with a small overshooting in the system for an object with an unfavorable combination of its model parameters.

13.6. Numerical optimization of the regulator for an object of dimension 3 × 3

Many publications have been devoted to the control of multi-channel objects. Numerical examples are given, which are most often limited to 2 × 2 dimensions, and the given statements are often extend to an arbitrary dimension $N \times N$. It should be recognized that the "induction" of any statements on an arbitrary order requires the fulfillment of two conditions: a) it is necessary to show that for a certain value of N this assertion is true, b) it is also necessary to show that if this assertion is true for N, then it is also true for $N + 1$, where N is an arbitrary integer. If the second is not proved, the distribution of previously obtained results to higher-order results can not be considered justified.

In this regard, even if there are results for objects of dimension 2 × 2, no less urgent is the study of problems with higher dimensionality, in particular, 3 × 3.

In addition, if in the transfer function of the object there are also links of pure delay along with the minimum-phase links, even for the 2 × 2 dimension the problem can not be solved by analytical methods, but can be solved by numerical optimization in mathematical modeling (simulation). However, in this case, the increase in dimension increases the complexity of solving the problem in a square, that is, in the transition from $N = 2$ to $N = 3$, the complexity of the problem increases in proportion to the ratio of the squares, namely: in 9/4 = 2.25 *times*.

In this case, even with the availability of powerful software and hardware, the minimization of the elements in the model becomes relevant. An important question is the validity of each element in the objective function and the validity of each element in the regulator.

In this chapter, the validity and importance of each such element is determined by the method of numerical optimization in mathematical modeling using the example of a three-channel control of object that contains in each channel a minimum phase link and a delay link in series with it.

Let discuss an object that has 3 inputs and 3 outputs, the elements of the matrix transfer function are third order filters with delay. An object can be described by a transfer function.

The transfer function of the object has the form:

$$W_S(s)=\begin{bmatrix} w_{11}(s) & ... & w_{N1}(s) \\ ... & ... & ... \\ w_{1N}(s) & ... & w_{NN}(s) \end{bmatrix}. \tag{13.11}$$

It is necessary to find the transfer function of the serial regulator, which would provide control according to the traditional requirements. Namely: autonomous control in static mode (that is, zero static errors for each channel), whenever possible minimal dynamic errors (small influence of control signals on all side channels), as small as possible overshooting (no more than 20%, and if it is possible, not more than 5%).

In general, the transfer function of the regulator can be described in the following form:

$$W_R(s)=\begin{bmatrix} q_{11}(s) & ... & q_{N1}(s) \\ ... & ... & ... \\ q_{1N}(s) & ... & q_{NN}(s) \end{bmatrix}. \tag{13.12}$$

We can put these requirements into the objective function all in its entirety, that is, if at least one of these requirements is not met, then the objective function will increase dramatically.

However, there is a simpler method, namely: the objective function can be constructed only on the basis of the integral of the sum of the error modules.

In a more complex form, terms can be introduced into the objective function, which increase under the following conditions: a) the overshooting exceeds a certain threshold; b) the product of the error by its derivative exceeds zero or a certain small positive threshold; c) the integral of the above specified value exceeds a certain threshold, and so on.

Earlier we proposed to introduce the "incorrect motion detector" (error growth detector), which calculates the integral of the positive part of the error product with its derivative. In the case of a multichannel object, it is necessary to take the integral of the sum of such products for each channel. When optimizing, we can use a comparison of the results with two cost functions: a) on the basis of the integral of the sum of the errors; b) on the same basis, but with the introduction of an incorrect motion detector.

The simplest regulator is the diagonal regulator, which is matrix transfer function (13.12) with non-zero elements only in the main diagonal. If this is not enough, it will be necessary to introduce nonzero terms into all elements of this matrix transfer function.

The easiest to control are objects in which the transfer functions in the main diagonal are greater than in the other elements. If this is not the case, but if this can be achieved by changing the numbering of the inputs or outputs, we recommend this. If this can not be achieved, we have to work with what is available. Therefore, examples of successful combinations of object parameters and unsuccessful combinations will be considered.

We recommend using the *VisSim*, since this software is created specifically for simulating dynamic feedback systems and for optimizing the regulators for them, although this software is not free from some drawbacks.

Example 49. Let consider an object with a transfer function (13.11), where $w_{ij}(s) = k_{ij}\ \exp\{-\tau_{ij}s\}\ /\ (a_{ij}\ s^2 + b_{ij}s + 1)$. Specific numerical values of the coefficients are given in *Table* 1.

Table 1

The object model coefficients

i	j	k_{ij}	τ_{ij}	a_{ij}	b_{ij}
1	1	5	2	1	1
1	2	3	1	1	2
1	3	2	3	1	3
2	1	4	1	2	4
2	2	4	2	3	6
2	3	2	2,5	2	8
3	1	4	1,5	1	5
3	2	3	1	2	3
3	3	2	2	3	3

It is required to find the regulator transfer function in the form (13.12).

In order to solve the task, we create a system project in the *VisSim*, as shown in Fig. 12.43–12.45. Since the circuit is too large, it is shown in separate fragments. During optimization, unite steps were applied to the input of the system, with the first input fed with zero delay, the second input with a delay of 40 *s*, the third input with a delay of 80 *s*. This makes the input signals linearly independent, which makes it possible to provide autonomous control with a regulator obtained during optimization. If this is not done, the result may not provide autonomy requirements.

The obtained equation of the regulator has the form:

$$W_{R1}(s) =$$

$$= \begin{bmatrix} -0{,}38+0{,}048/s & 0 & 0 \\ 0 & 1{,}125-0{,}0247/s & 0 \\ 0 & 0 & 0{,}473+0{,}098/s \end{bmatrix}. \quad (12.13)$$

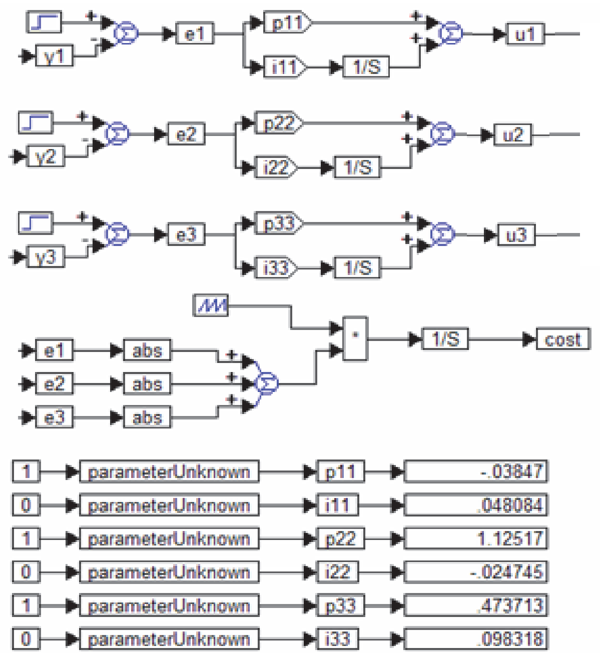


Fig. 13.43. The structure of the diagonal PI regulator when trying to use only the proportional and integrating channels and use only the diagonal elements in the regulator matrix (part 1 is the regulator)

Figure 13.45 shows the obtained transient processes in a system with such a diagonal regulator. From these processes it is clear that in the second channel the process goes on the wrong side in the first third of the graph, namely: with the passage of time the output value is removed from the prescribed value.

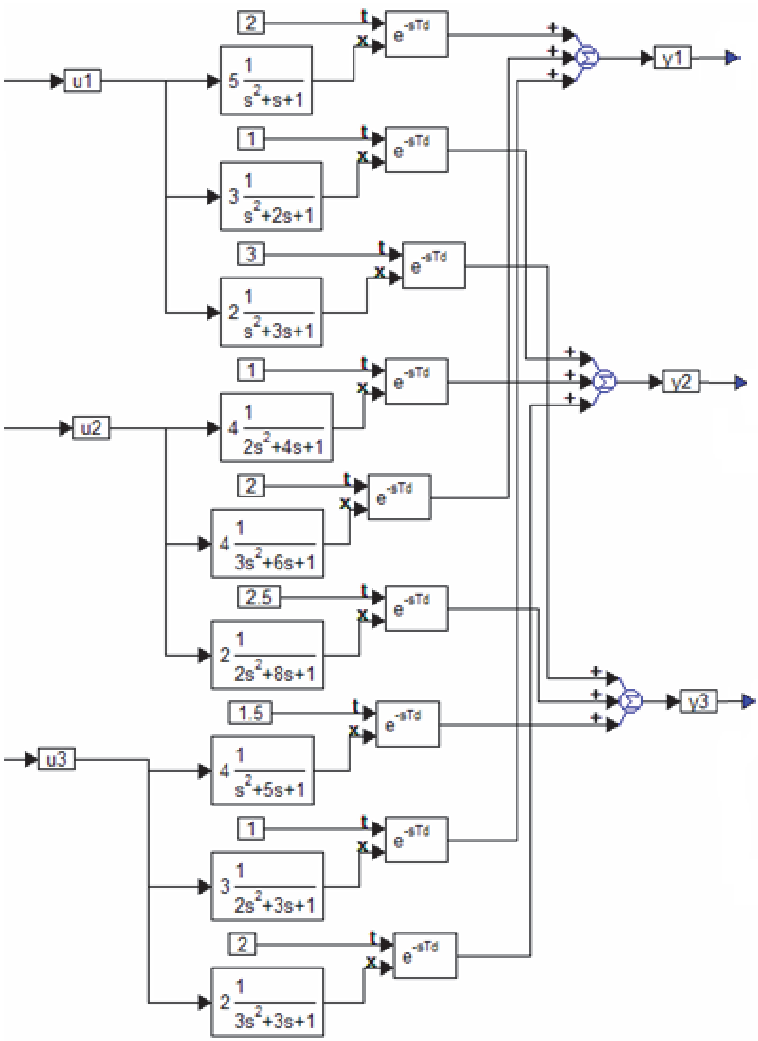


Fig. 13.44. The structure of the diagonal PI regulator when trying to use only the proportional and integrating channels and use with only the diagonal elements in the regulator matrix (part 2 is object)

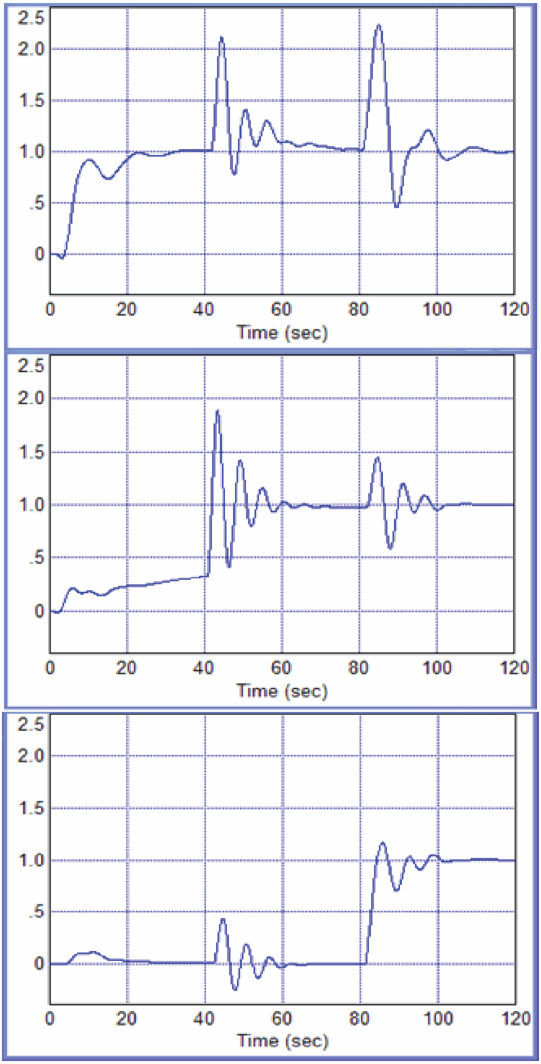


Fig. 13.45. The structure of the diagonal PI regulator when trying to use only the proportional and integrating channels and use only the diagonal elements in the regulator matrix (part 3 is oscilloscopes with output signals: first, second and third channels)

This behavior of the process is explained by the negative coefficient before the integral component of the regulator of the second channel, that is, the polynomial at the intersection of the second row and the second column. For a given object structure (all elements of the main diagonal are positive), the coefficient in the integrating path must be positive.

Thus, the received regulator should be recognized as not corresponding to the task in view.

Example 50. To control the same object, we introduce derivation into regulator. Then the regulator will be as shown in Fig. 13.46. The resulting transients are shown in Fig. 13.47.

Now there is no any wrong on statics fragment in any channel. But overshooting is great. The received regulator is described by the following transfer function:

$$W_{R2}(s) =$$

$$= \begin{bmatrix} 0.123 + \dfrac{0.203}{s} + 0.127s & 0 & 0 \\ 0 & 0.945 + \dfrac{0.926}{s} + 0.663s & 0 \\ 0 & 0 & 0.228 + \dfrac{0.266}{s} + 0.297s \end{bmatrix}. \quad (13.14)$$

Now all coefficients for integrators are positive. The static error in each channel is zero, which is expressed in the fact that all transient processes terminate with time on those values that are fed to the input of the system. The overshooting in the first channel reaches 150%. In the second channel, it is only slightly less – about 110%, in the third channel it reaches 60%.

Thus, with the considered regulator, the problem as a whole is solved, but the overshooting is extremely large, so the result can not be considered satisfactory either.

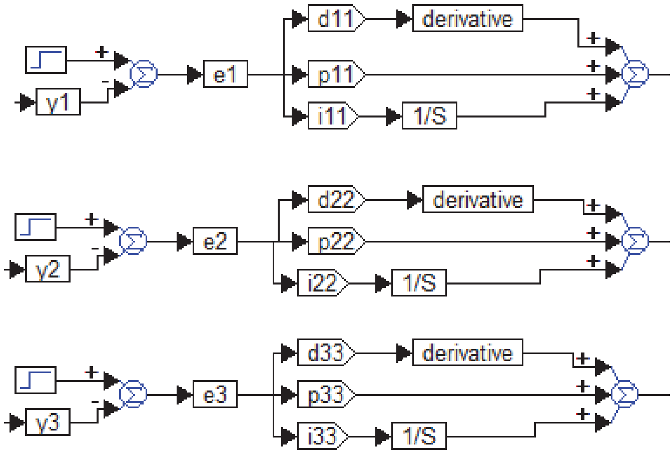


Fig. 13.46. The diagonal PID regulator in accordance with Example 50

Example 51. Let consider the same object, we will use a regulator which in the main diagonal of the matrix of the transfer function contains scalar PI regulators, and in its other elements it contains proportional regulators. In doing so, we will use the same cost function and the same input actions. Figure 13.48 shows the corresponding structure of the regulator, and Fig. 13.49 presents the results in the form of transient processes. The transfer function of the received regulator has the form:

$$W_{R3}(s) =$$

$$= \begin{bmatrix} 0.063 + \dfrac{0.053}{s} & 0.064 & -0.007 \\ -0.0532 & 0.393 + \dfrac{0.133}{s} & 0.064 \\ 0.017 & -0.33 & 0.269 + \dfrac{0.0685}{s} \end{bmatrix}. \qquad (13.15)$$

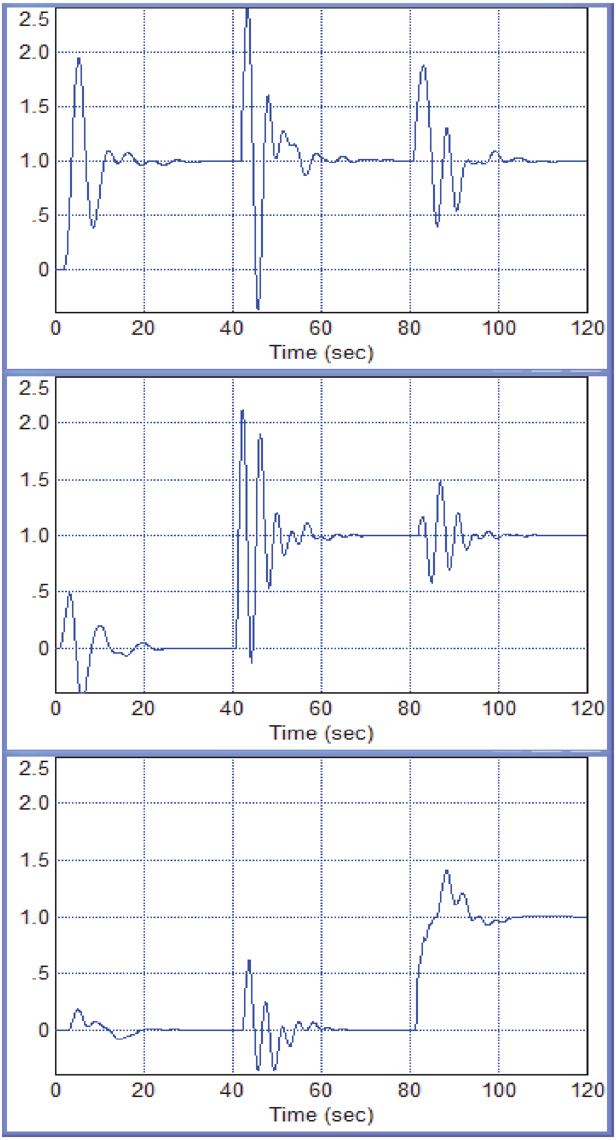


Fig. 13.47. The results of optimizing the diagonal PID regulator of Example 50

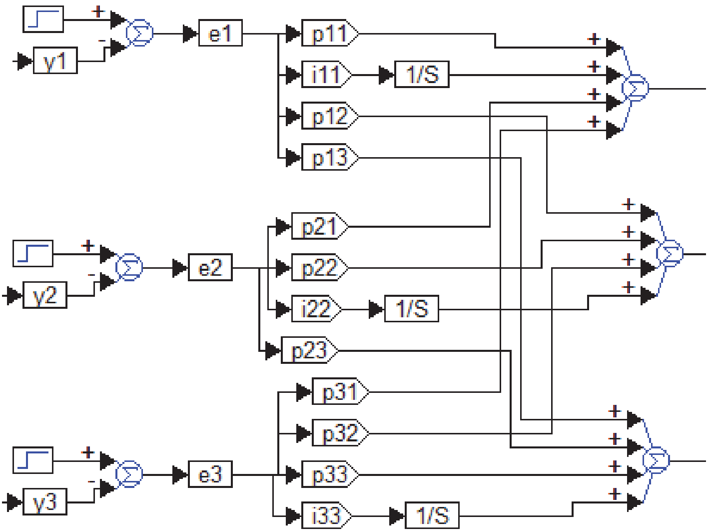


Fig. 13.48. Regulator of Example 51

The results are much better than in the previous examples, however, on the first channel the overshooting is still large, about 80%. Such a system may be acceptable for some applications, but in most cases such a large amount of overshooting still does not satisfy the requirements of the technological process.

Example 52. Let consider the same object and the same regulator, but in doing so, we introduce error growth detector into the cost function. Only positive parts are taken from these products, which are summed up with integration, after which the result is added to the value function. Since the calculator of the value function already contains an integrator, we can confine ourselves to only one common integrator, and the summation is performed at its input. Figure 13.50 shows the corresponding structure for calculating the value function. A weighting factor of ten is used.

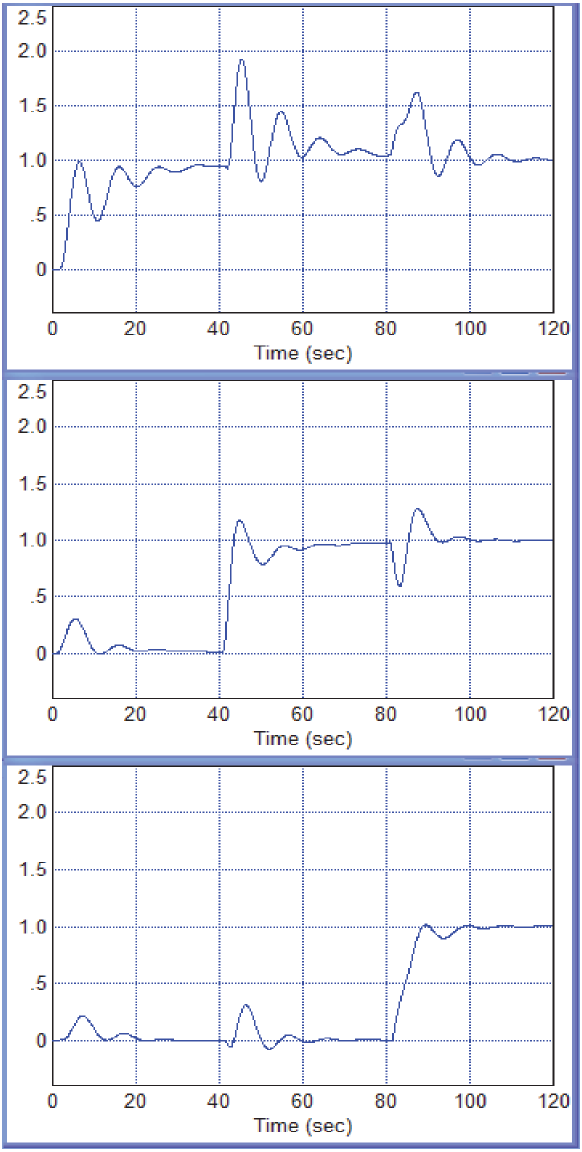


Fig. 13.49. Results of optimization of the regulator in Example 51

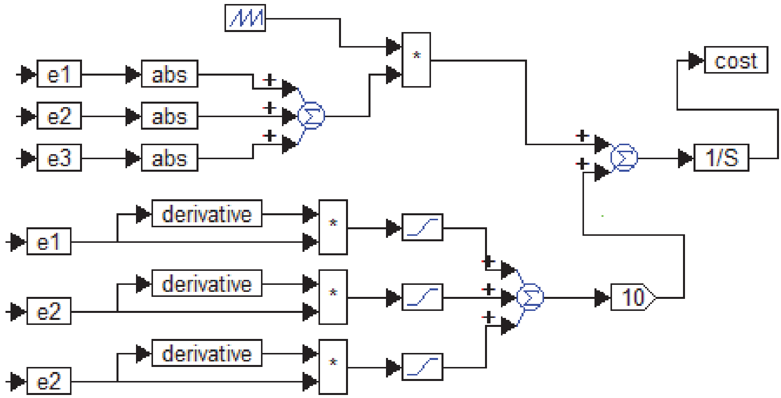


Fig. 13.50. Structure for calculating the cost function, including error growth detector

The received regulator has the following transfer function:

$$W_{R3}(s) =$$

$$= \begin{bmatrix} 0.068 + 0.051/s & 0.043 & 0.01 \\ -0.157 & 0.163 + 0.068/s & -0.075 \\ 0.048 & -0.082 & 0.05 + 0.038/s \end{bmatrix}. \quad (13.16)$$

The resulting transients are shown in Fig. 13.51. Overshooting in the first channel now does not exceed 50%, and in other channels no more than 25%.

We can use a different weighting factor, for example, equal to five. The resulting transient processes are shown in Fig. 13.52. It can be seen that the overshooting in the first channel increased to 60%. Therefore, this result is not better than the result with the regulator (6).

We can also increase the weighting factor, for example, to twenty. The corresponding transient processes are shown in Fig. 13.53. The overshooting in the first channel dropped to 40%, but the duration of the transient processes increased greatly, they became tightened. Therefore, the result with the regulator by the relation (13.16) should be recognized as the best with such a predetermined forward structure among all obtained in this example and in the previous examples.

Example 53. Let consider the same object and the same objective function as in Example 52, but we will use a regulator in which in the main diagonal the scalar PID regulators are contained, and the other elements of the matrix will are only proportional links. We also used weighting factors equal to five, ten and twenty. The resulting transient processes are shown in Fig. 13.54, 13.55 and 13.56, respectively. On the processes shown in Fig. 13.54, overshooting of the first channel is not more than 40%, in other channels is much smaller. However, the processes are not too tight. On other graphs, the processes are not better; there is a tightening of transient processes. Therefore, it is suggested to prefer the result obtained with a weighting factor of five.

The received regulator is described by the following transfer function:

$$W_{R4}(s) =$$

$$= \begin{bmatrix} 0.087 + \dfrac{0.0787}{s} + 0.059s & -0,012 & -0,0007 \\ -0,117 & 0.3 + \dfrac{0.156}{s} + 0.24s & -0.137 \\ 0,0016 & -0.0414 & 0.199 + \dfrac{0.064}{s} + 0.095s \end{bmatrix} . \quad (13.17)$$

For comparison, Fig. 13.57 shows transient processes with the same regulator calculated using the zero coefficient for the motion detector. In this case, the overshooting in the first channel is 110%, which shows that even with the most complex structure of the regulator, the rejection of the error growth detector leads to the fact that the task is not so successfully solved.

Thus, it is shown that the use of the error growth detector was one of the key approaches necessary to solve the problem of design of regulator for controlling of the three-channel object.

Other important principles of optimization are that the input effects must be linearly independent, the cost function includes the integral of the sum of the error modules and the most complex (and therefore most effective) PID regulators should be used in the main diagonal. In this case, only proportional regulators can be used in the non-main connections, and at the same time control can be obtained acceptable, namely: autonomous control is provided, zero static errors for each channel, overshoot does not exceed 40% in the worst case.

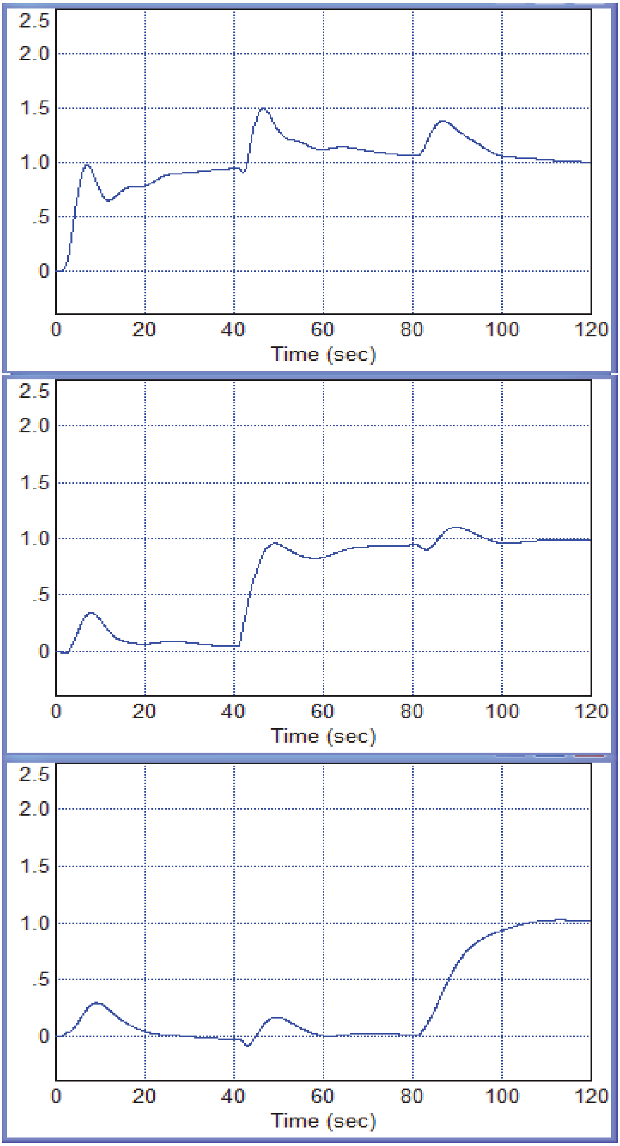


Fig. 13.51. The resulting transient processes from Example 52

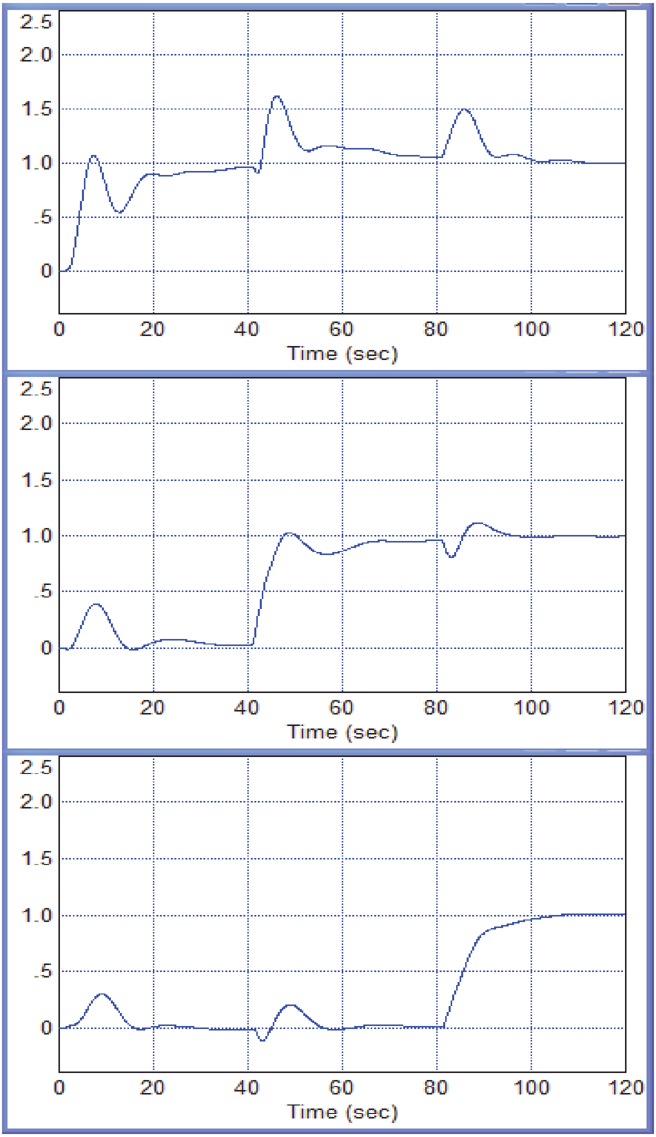


Fig. 13.52. The resulting transient processes of Example 52 using a weighting factor equal to 5

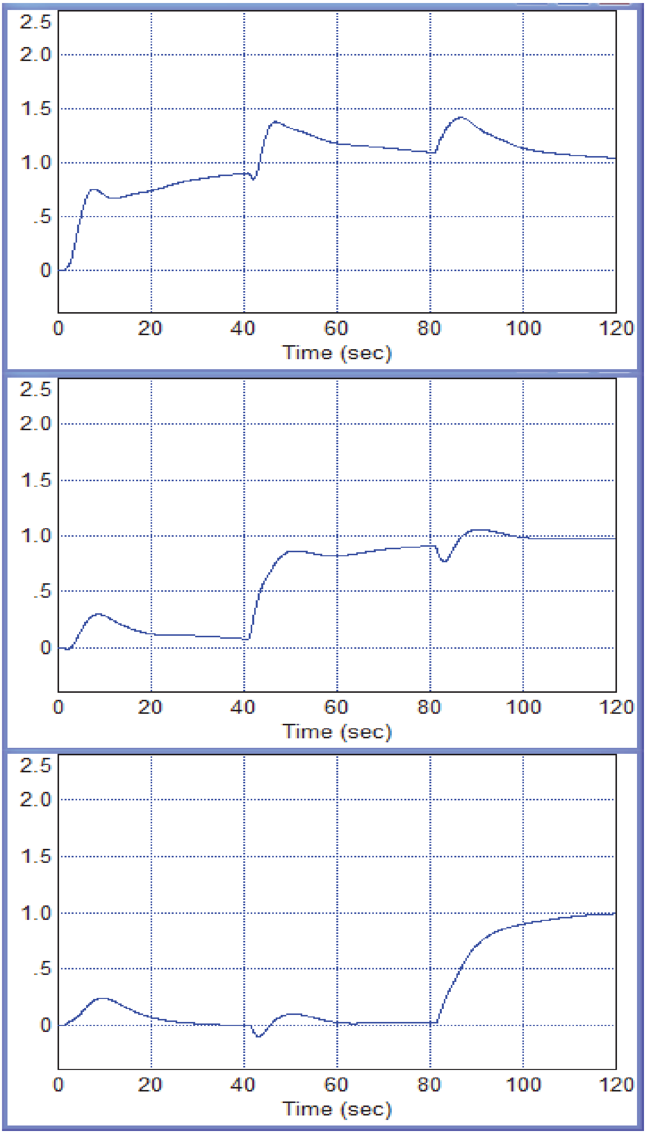


Fig. 13.53. The resulting transient processes from Example 52 using a weighting factor equal to 20

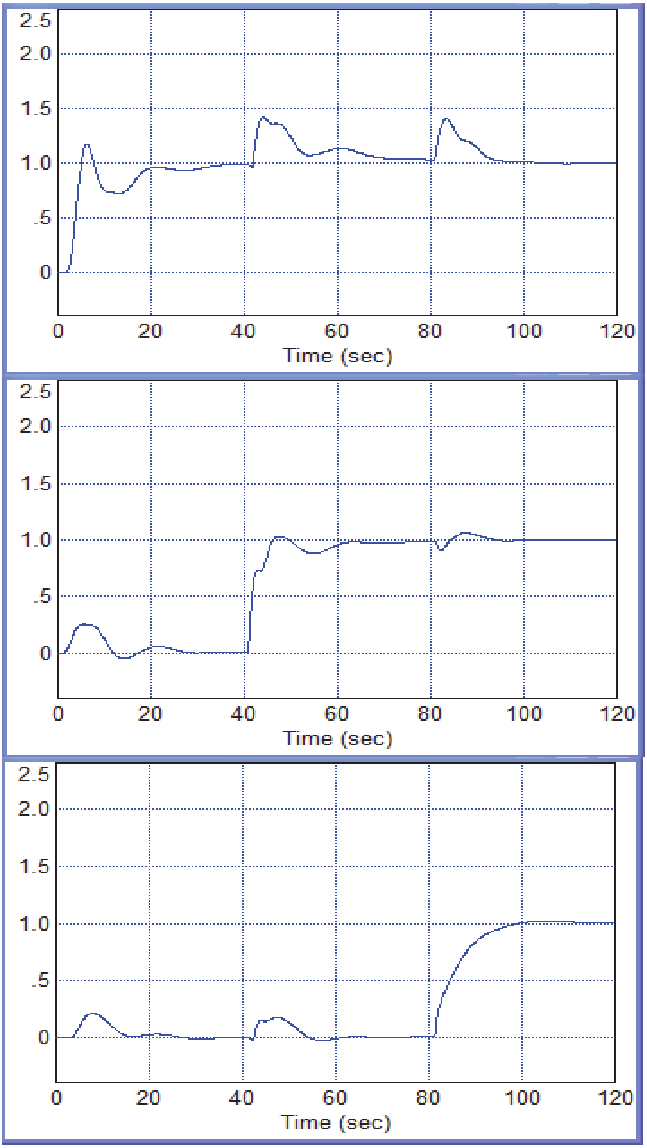


Fig. 13.54. The resulting transient processes of Example 53 using the weighting factor equal to 5

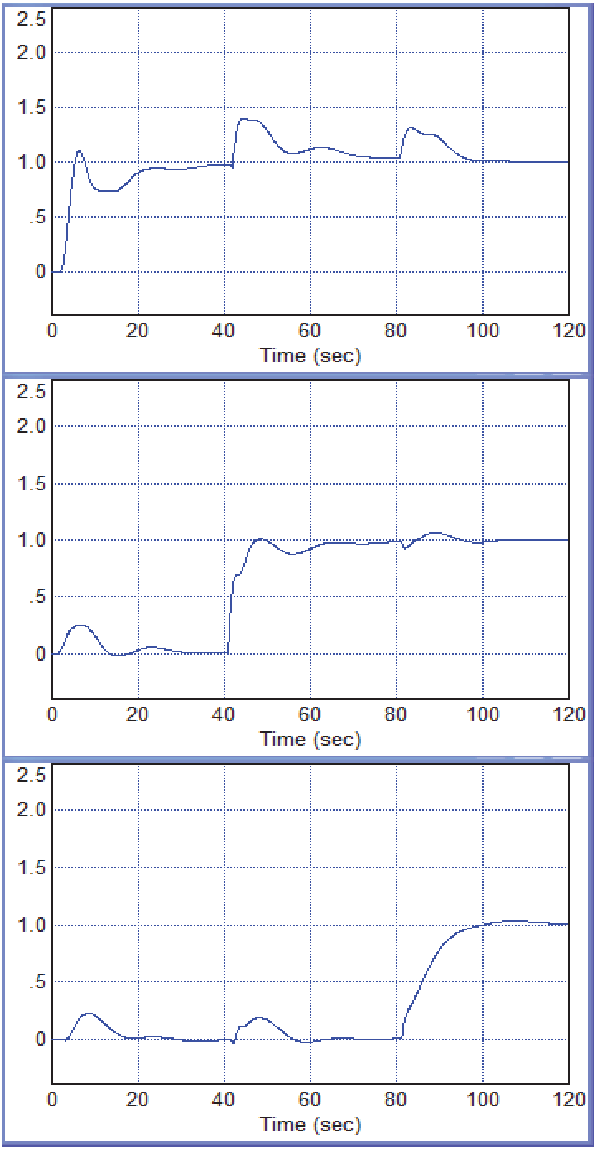


Fig. 13.55. The resulting transient processes of Example 53 using a weighting factor equal to 10

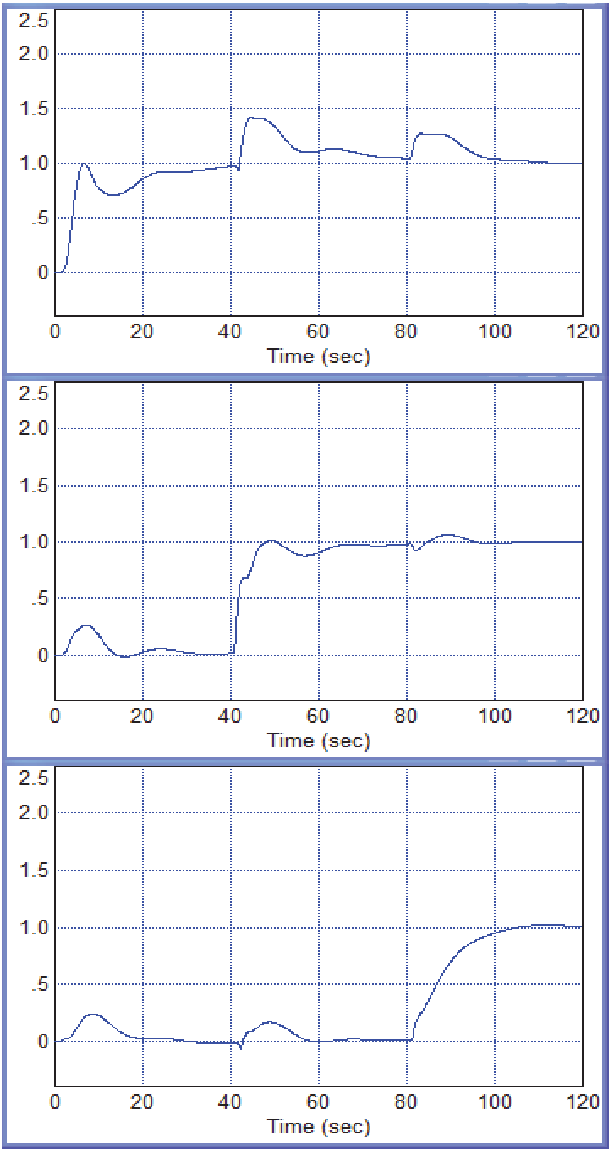


Fig. 13.56. The resulting transient processes of Example 53 using a weighting factor equal to 20

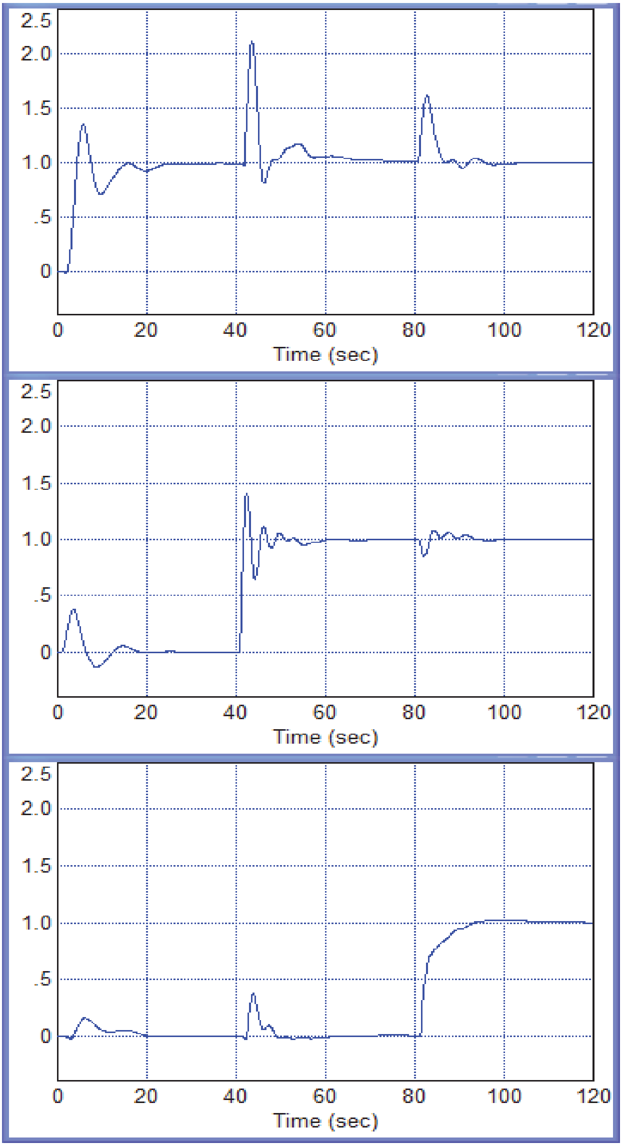


Fig. 13.57. The resulting transients in Example 53 when using a weighting factor equal to 0 for comparison

Naturally, more complex regulators can give better results; however, the use of PID regulator in each channel would require optimization of twenty-seven coefficients, which exceeds the capabilities of the version of *VisSim* that we are using.

However, if the conditions of the problem still require a reduction in overshooting, one of the following ways can be proposed.

First, we can suggest using the latest version of *VisSim* or other software to optimize the required number of parameters.

Secondly, if the first option is not available, we can suggest using, for example, the method of frozen coefficients. Namely, after optimizing the largest possible number of parameters, a developer can fix them, and then optimize the next part of the parameters, then fix them, and then optimize the remaining parameters. After that, he can return to the first group of parameters and so on until an acceptable quality control will be obtained.

14. INVESTIGATION OF THE NUMERICAL OPTIMIZATION TOOLKIT FOR CONTROL OF AN OSCILLATORY UNSTABLE OBJECT

14. 1. Statement of the problem

Let us consider a linear oscillatory unstable object described by a transfer function expressed as a Laplace transform of the following form:

$$W(s) = \frac{-s+1}{s^3 + 14s^2 - 4s + 50} . \qquad (14.1)$$

The objective is to find a regulator, which provides stable control with the best possible quality of the transient process.

Traditionally, in automated systems, is it most common to use a serial PID regulator, i.e. a regulator having proportional, integral and derivative links that are combined in parallel. The regulator is switched on at the input of the object in series with it, while the setpoint for the system is fed to the first input of the regulator, and the output signal of the object is fed with a single negative coefficient to the second input of this regulator. Two controller inputs are provided by using the signal adder.

The transfer function of the PID-regulators is the following:

$$W_{PID}(s) = k_P + \frac{k_I}{s} + k_D s . \qquad (14.2)$$

Here k_P, k_I, k_D are the coefficients to be optimized.

For the PIDD-regulator, the transfer function is:

$$W_{PIDD}(s) = k_P + \frac{k_I}{s} + k_D s + k_{DD} s^2 . \qquad (14.3)$$

Here, one more coefficient k_{DD} is added for the path with double derivation.

Thus, the goal is to find the coefficients for regulators (14.2) or (14.3), or to suggest a different structure of the regulator and calculate the coefficients for this structure.

The solution must be sufficiently robust, i.e. not to change significantly in the case of insignificant changes in the operating conditions of the system or the coefficients of the regulator or the object.

In some publications this requirement is not explicitly formulated, which is a mistake, because without this requirement the regulator cannot be used in practice.

14.2. Methods for solving the problem, theory and practical results

14.2.1. Optimizing the PID-regulator

The most obvious method for solving the problem is to optimize the PID-regulator (14.2) numerically. Fig. 14.1 shows the block diagram for modeling the system and its optimization in the software *VisSim*. The cost function consists of an integral of two terms:

$$\Psi(T) = \int_0^T [\psi_1(t) + \psi_2(t)]dt . \qquad (14.4)$$

Here

$$\psi_1(t) = | e(t) | dt . \qquad (14.5)$$

$$\psi_2(t) = k_w \max\{0, f(t)\} . \qquad (14.6)$$

$$f(t) = e(t)\frac{de(t)}{dt} . \qquad (14.7)$$

The weight coefficient k_w is chosen empirically to be 10. When k_w is smaller than 10, the second term in (4) does not significantly affect the result, while when k_w is larger than 10, it drecreases the quality of the procees. The values obtained for the regulator coefficients are displayed in the lower right corner of Fig. 14.1. Fig. 14.2 shows the resulting transient process in the system. The input signal in this case is a single step function at $t = 0.1$ *s*. The delay of 0.1s from the beginning of the process is necessary for the robustness of the solution.

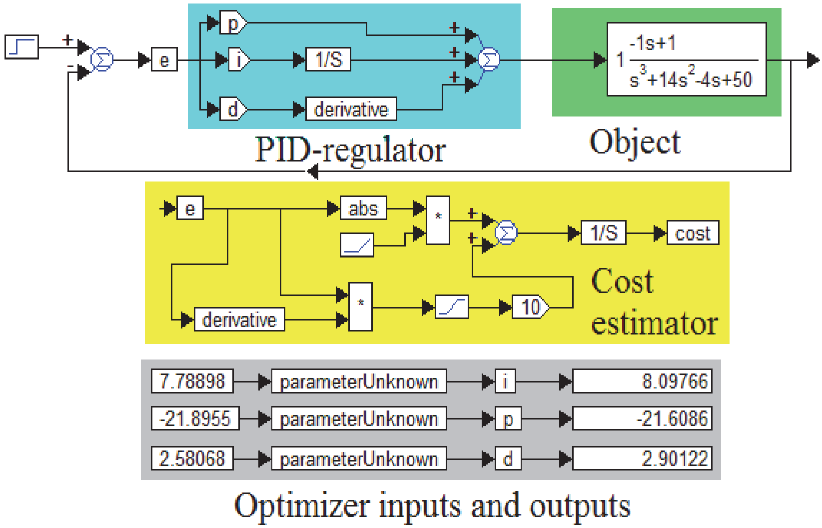


Fig. 14.1. Structural diagram of a system with a PID-regulator and the result of finding the coefficients by the method of numerical optimization (the values of the coefficients are shown in the lower right corner)

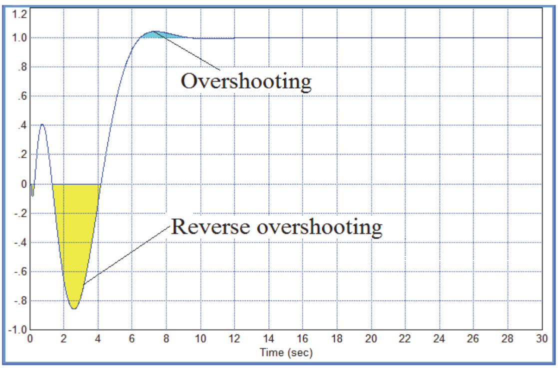


Fig. 14.2. Transient process in the system with PID-regulator

The duration of the obtained transient process is about 10 *s*. The overshoot does not exceed 5%, but there is a reverse overshoot, which is about 85% of the value of the input step function.

Although formally the system with reverse overshoot is stable, it is highly undesirable that instead of moving in the prescribed direction, the system moves in the opposite direction at the beginning of the transient process. Certainly, such a move should be eliminated, and if this is not possible, it must be minimized.

14.2.2. The optimization of PIDD-regulator

Since adding a double derivative term gives additional possibilities of controlling the object, we try to solve the problem by numerical optimization of the PIDD-regulator (14.3). Fig. 14.3 shows the block diagram for modeling the system and its optimization in the sortware *VisSim*. The cost function and the input are the same as above.

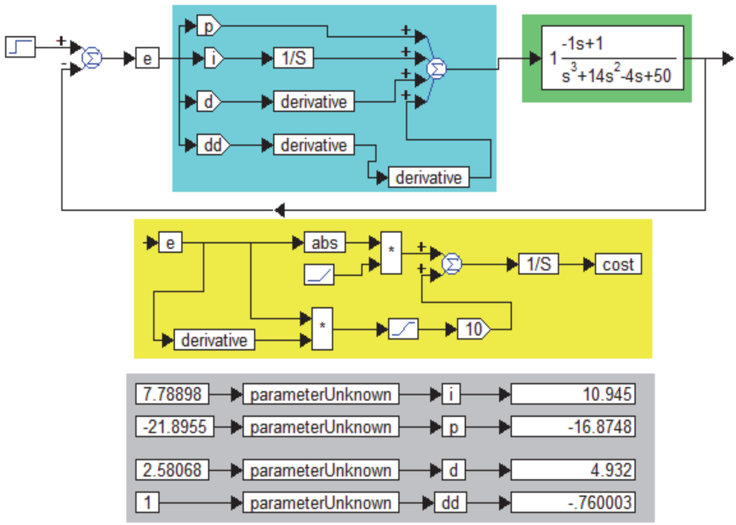


Fig. 14.3. Structural diagram of the system with PIDD-regulator and the result of finding the coefficients by the method of numerical optimization (the values of the coefficients are shown in the lower right corner)

Fig. 14.4 shows the transient process in the resulting system. It can be seen that the duration of the transient process was reduced to 8 *s*, the overshoot was reduced, the reverse overshoot also decreased to 70%. The system has improved, however, its performance is still not satisfactory.

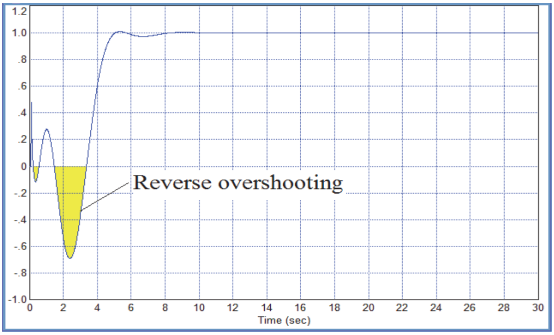


Fig. 14.4. Transient process in the system with PIDD-regulator

The reverse overshoot now occurs twice, first by 10%, then by 70%. The first overshoot can be neglected in comparison with the second one, therefore, the quality of the system has to be assessed given the second overshot.

14.2.3. Bypass channel application

The Bypass channel is a development of the idea of the Smith's predictor. It is connected in parallel to the object and has a transfer function that is negligible in the low-frequency region, and in the high-frequency region it is larger than the transfer function of the object and at the same time it has better properties (it fades out more smoothly).

The transfer function of the Bypass channel can be formed using the following considerations.

If all the coefficients of the numerator of the transfer function of the object are positive, the control is much easier.

Therefore, the desired transfer function of the object can be written in the following form:

$$W_D(s) = \frac{s+1}{s^3 + 14s^2 - 4s + 50}. \tag{14.8}$$

Then subtracting (14.1) obtains the transfer function of the Bypass channel from (14.8):

$$W_{PB}(s) = \frac{2s+1}{s^3 + 14s^2 - 4s + 50}. \tag{14.9}$$

Fig. 14.5 shows a block diagram for optimizing a system with the Bypass channel.

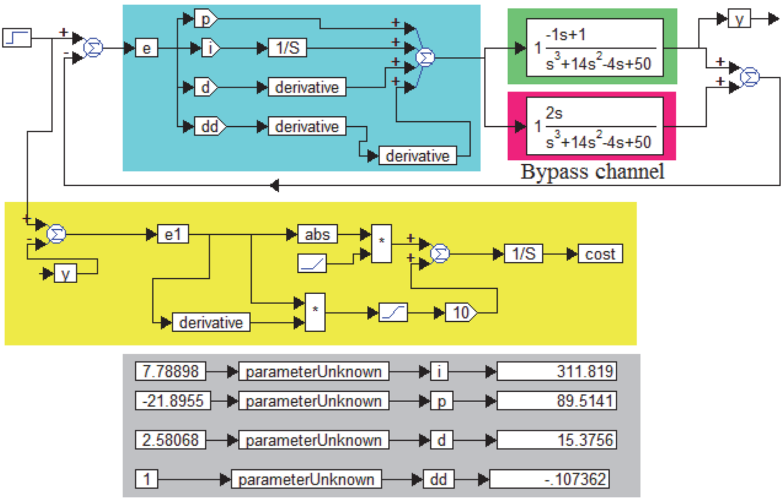


Fig. 14.5. The block diagram with the Bypass channel and the optimization result in the form of the PIDD-regulator coefficients (coefficients are shown in the lower right corner)

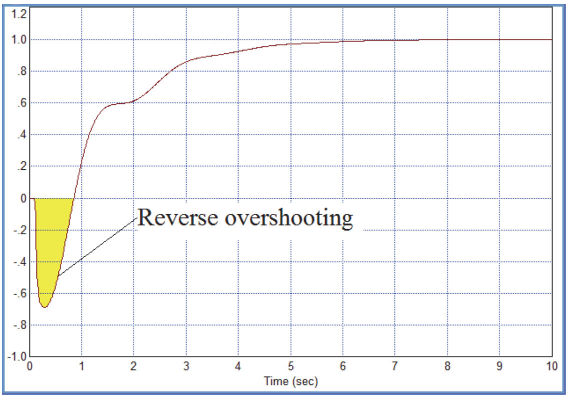


Fig. 14.6. Transient process in a system with the Bypass channel

As can be seen, the transient process has been somewhat improved. The reverse overshoot remains problematic: although its maximum value did not change, it is still 70%. The duration of the process is about 7 *s*. Compared to the result of using a PIDD regulator, the response has improved since the duration of the reverse overshoot was reduced from 2 *s* to 1 *s*. However, this improvement is not too significant.

14.2.4. Using pseudo-local links and an open object model

The method of pseudolocal loops is based on the method of local loops.

If the object had outputs of intermediate signals, these signals could be used to further stabilize it. Such loops are called local.

Since the real object does not have intermediate outputs, a developer can try to simulate what the signals at such intermediate points would look like by using the mathematical model of the object with the same input signals as the ones given to the real object.

If the model of the object is completely identical to the real object and does not have intermediate outputs, such a model is not suitable for creating pseudo-local links, since this model is closed. An open model is a model in which all intermediate signals are available for their use. Fig. 14.7 shows two models of the same object.

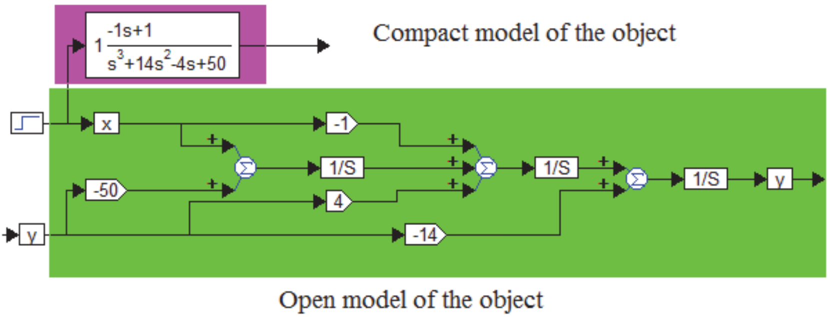


Fig. 14.7. The model of the object in the form of a transfer function (above) and in the form of an open model (below)

The upper model is closed, because intermediate signals are not available for use. The lower model is open, because there are intermediate signals available. Additional inputs are also available for controlling the system. The signals on the outputs of both models must completely coincide, since these structures are described by an identical mathematical description.

Since in the simulation with the software *VisSim* calculations are carried out by steps, these output signals may be slightly different. In particular, Fig. 14.8 shows both output signals. It can be seen that the signal at the output of the closed model differs somewhat in amplitude (line 2), but in general it is close enough to the signal at the output of the open model (line 1).

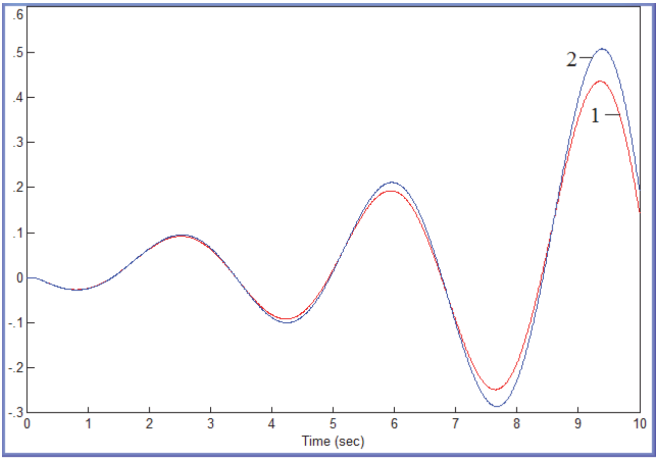


Fig. 14.8. Graphs of transient processes in the initial model of the object (line 1) and in the open model of the object (line 2)

Theoretically, using signals from the output of an open model, by setting them on the input of an object, one can hope that as a result the mathematical model of an object will change in the same way that its model would change if these signals were taken directly from the object.

In order to calculate which feedbacks are necessary, we first need to find the model closest to the actual model of the object. It should be such that the task of managing such an object was simple, and the result was close to what was desired.

The following form of the desired transfer function of the object is found:

$$W_D(s) = \frac{s+1}{s^3 + 14s^2 + 40s + 50}\,. \tag{14.10}$$

In this ratio, positive ones replace all negative coefficients, and the coefficient of the second power of the argument s in the denominator is increased tenfold. Such transformations make the transfer function stable and the corresponding object is easily controllable. For this transfer function, we can also make an open model. The corresponding open model is shown in Fig. 14.9 below. We can also add local links to the original open model to convert it to the desired one. The corresponding structure is shown in Fig. 14.9 above. Formally, these structures correspond to the same mathematical model, but in practice, they can work differently for the reasons indicated above.

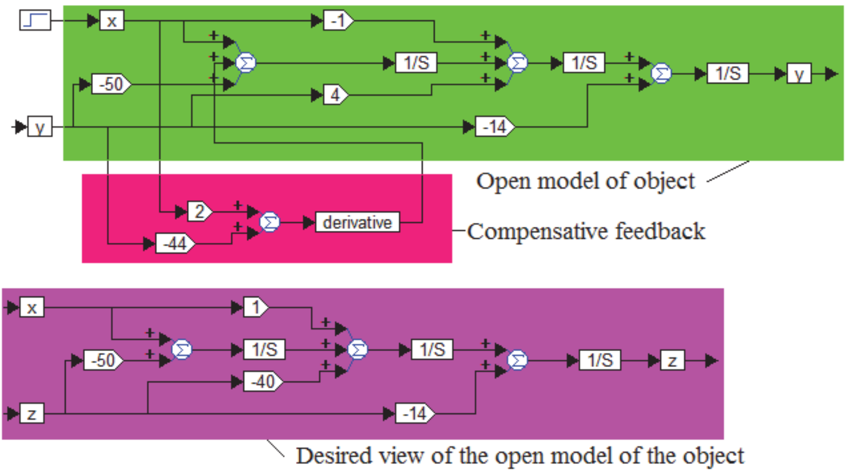


Fig. 14.9. Structures for modeling the improved model of the object in the form of a modified model (below) and in the form of an open model with compensating links (at the top)

Fig. 14.10 shows the transient processes at the outputs of the structures shown in Fig. 14.9. It can be seen that these signals do not coincide. In this case, the input signal is a single step function generated at time $t = 0$. If a delay a delay of at least the step size of the sampling, i.e. $t = 0.1$ *s*, is introduced, then the graphs of the transient processes at the outputs of these structures almost completely coincide, as shown in Fig. 14.11. It is for this reason that in the previous sections the input signal was used with a delay for a specified time interval.

Based on this, pseudo-local links using an open model could work efficiently, but this requires additional research.

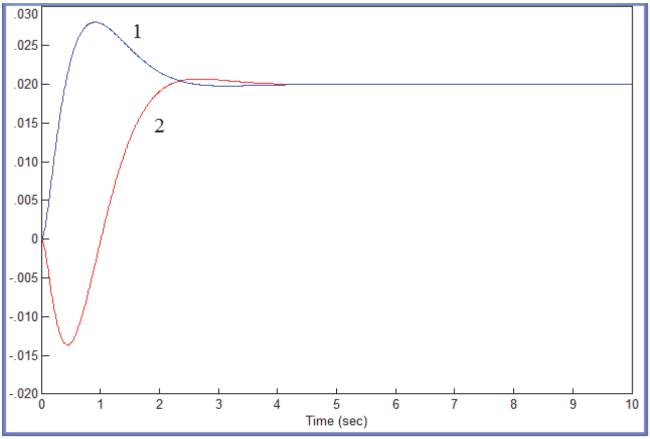


Fig. 14.10. Graphs of the improved model of the object in the form of a modified open model (line 1) and in the form of an open model with compensating links (line 2); the both graphs are obtained without delay in the input effect, therefore the graphs differ dramatically

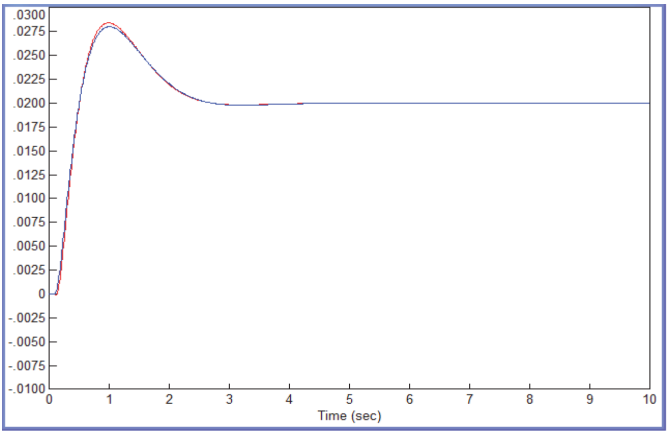


Fig. 14.11. Graphs of the improved model of the object in the form of a modified open model in the form of an open model with compensating links after the introduction of the delay of the input action for a time of 0.1 *s*; the graphs practically coincide

Fig. 14.12 shows the structure of a system with pseudo-local loops based on the open model of the object. Fig. 14.13 shows the transient process obtained at the output of this structure. The transition process takes only 30 seconds to the value $6 \cdot 10^{-82}$, i.e. practically in a minus infinity. Numerical optimization of the regulator for such an object does not give a result, because the process stops due to an unacceptable error.

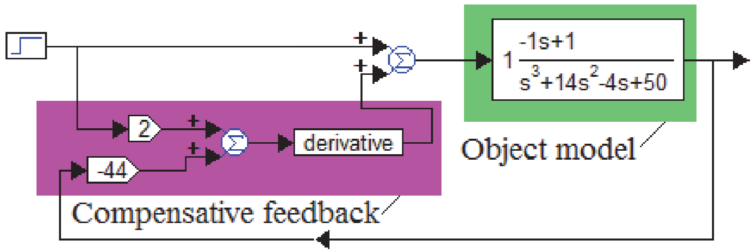


Fig. 14.12. Structure of the system with compensating pseudo-local feedbacks obtained by using the open object model

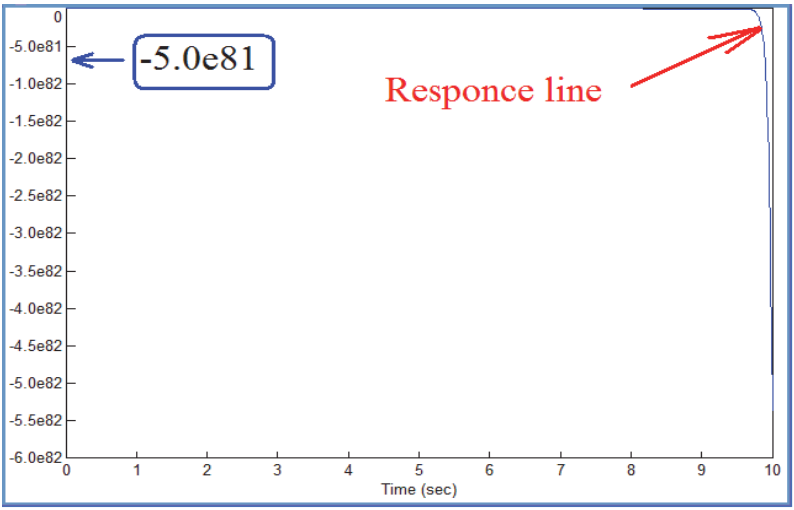


Fig. 14.13. An unstable transient process, obtained by modeling the system with compensating feedbacks shown in the previous figure

This approach can be further developed. For this purpose, the coefficients of local feedback in the structure of Fig. 14.12 can be designated with letters and the software should perform optimization by including these alphabetic variables in the number of optimized parameters. Fig. 14.14 shows the corresponding block diagram: in the lower right-hand corner are the numerical values for these coefficients and PI-regulator coefficients are given. Fig. 14.15 shows the obtained transient process in the system.

This process is stable but inferior to the processes discussed above, since the reverse overshoot is about 95%, and the typical overshoot of about 60% is still present.

On this basis, we can conclude that the method chosen did not yield a sufficiently effective result; other methods achieve better results.

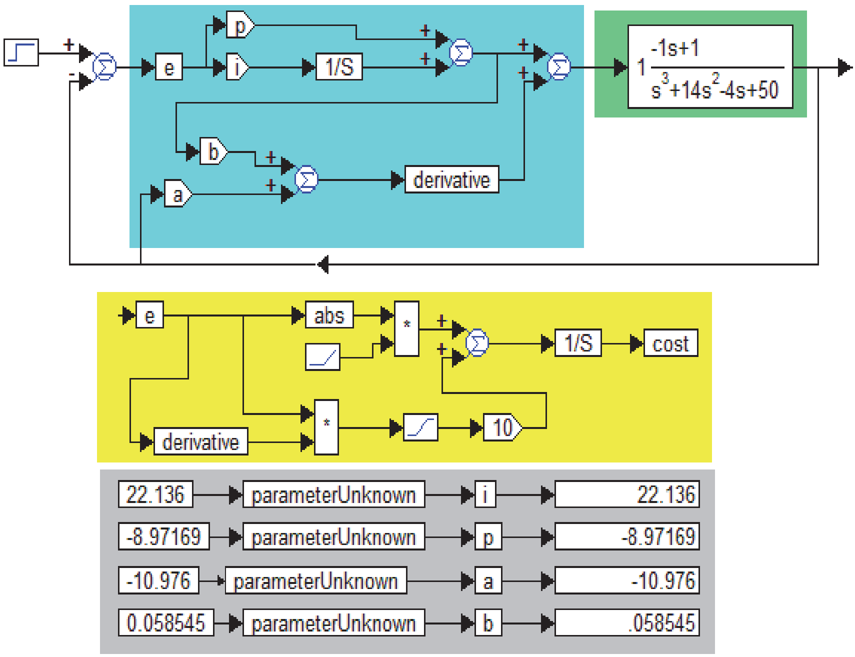


Fig. 14.14. Structural diagram of the system with PIDD-regulator and the result of finding the coefficients by the method of numerical optimization (the values of the coefficients on the lower right)

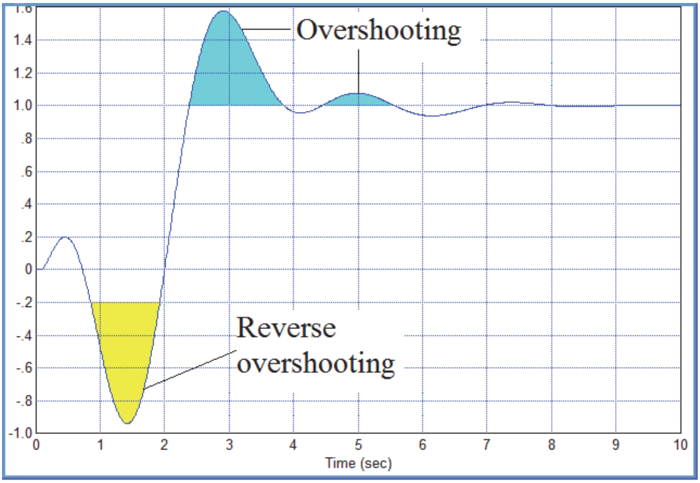


Fig. 14.15. Transient process in a system with PI-regulator and pseudolocal loops, the coefficients of which are calculated by the numerical optimization method

14.2.5. The danger of obtaining non-robust solution

Summarizing the use of the known methods of calculating the regulator by the use of numerical optimization, we note a possible source of error when using these methods. As indicated above, the introduction of a small input delay changes the simulation result significantly. Without such a delay, the output signals of the open and closed models of the object do not coincide, and when using such a delay, they are identical. This gives us a reason to prefer the use of this delay.

However, system developers would not necessarily try to use an open model and may not detect this difference in the simulation results.

In this paper, for the first time, we state that at least a minimal delay in such modeling and in the optimization of the regulator is extremely necessary. We suppose that developers do not use any delay in the input signal and calculate the regulator by numerical optimization. The structural scheme for this case is the same as the scheme, shown in ***Fig***. 14.3, but the result of its application is different.

The coefficients obtained have the following values: $k_I = 17.2$; $k_P = -2.24$; $k_D = 9.6$, $k_{DD} = 1.07$. The resulting transient process is shown in

Fig. 14.16. This process has undoubted advantages over the previous processes. Indeed, the reverse overshoot is only 40%, which is almost half of the best result obtained by other methods. There is almost no forward overshoot and the transient time is about 8 *s*. Therefore, such a result should be preferred compared to all the rest. But if applied in practice, the system will work quite differently, because this result is not robust. Small deviations in the simulation lead to significantly different results. If in this system we introduce a delay of at least one step of the discreteness of the simulation, i.e. on the value $t = 0.1$ *s*, the graph of the transient process will change significantly. If this value changes further, the type of the transient process will no longer change and will remain the same. For comparison, Fig. 14.17 shows two graphs with a delay t = 0.1 *s* and t = 4 *s*, the forms of these two graphs are absolutely identical: the same form could be seen as invariant with respect to the value of the delay of the input signal, zero pause is an exception, and in this case the graph is as shown in Fig. 14.16. Therefore, a zero delay cannot be used in the numerical optimization of the regulator. This is the reason why we strongly recommend using a non-zero delay in all cases: such a refinement has not been published before.

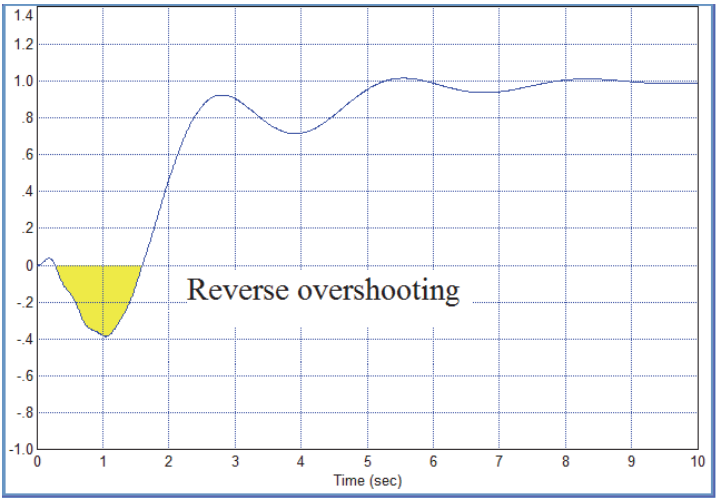


Fig. 14.16. Transient processes when the input signal is input immediately

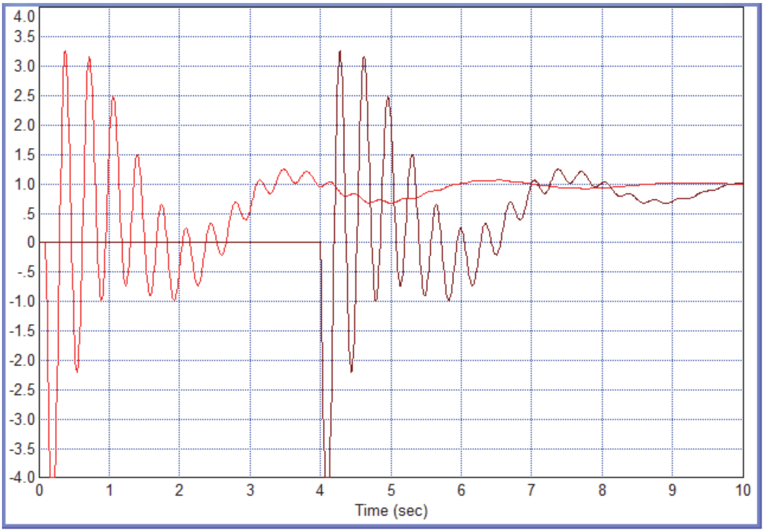


Fig. 14.17. Transient processes in the system for various delay values from 0.1 *s* to 4 *s*; it can be seen that the graphs are identical in the form

14.3. The proposed modification of the known methods

The best and at the same time not too complicated result can be considered the result with a sequential PID-regulator, shown in Fig. 14.4. The result with the PID- regulator shown in Fig. 14.4 is commensurable with this shown in Fig. 14.2. Both of these results are robust, so they are reliable. But both of these results are also characterized by a significant reverse overshoot, so they are not good enough for use in practice.

To improve them, the result of the system-controlled bu a PID regulator is treated as a new object for control. Such a system can be further covered by a single feedback: a sequential regulator, an integral or more complex regulator, such as the PID- regulator, can be used.

Fig. 14.18 shows a block diagram for modeling such a two-loop system and for optimizing its regulator. In this case, the composite block "Fast Loop" contains a system with feedback and a PID-regulator, the coefficients of which are fixed in accordance with the obtained numerical values from the structure of Fig. 14.3.

Unfortunately, the procedure of automatic optimization in this case is unstable and does not allow for obtaining a usefull result.

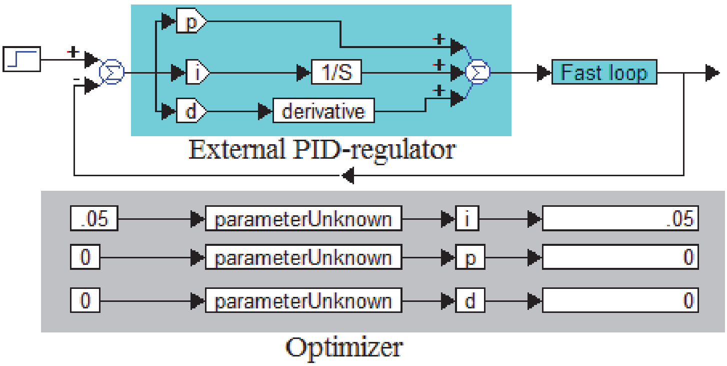


Fig. 14.18. Structural scheme with an integrator in the external circuit (using a PID-regulator with zero coefficients in the proportional and derivative links)

Therefore, we can use the known property of a system with an integrator and an object whose response is stable and corresponds to the response of the low-pass filter. Such a system is stable and is characterized by a zero static error if the coefficient of the integrator is sufficiently small. Empirically, putting the coefficients of the proportional and derivative links equal to zero, we obtain the approximately optimal result with a coefficient equal to 0.12. The processes in Fig. 14.19 illustrates that with a larger integrator factor overshooting occurs, and with a smaller coefficient, the process duration increases, while the selected value corresponds to a process in which the overshooting still remains insignificant and the process duration is sufficiently small in comparison with other possibilities. In this case, the reverse overshooting still remains, but its value does not exceed 5%, and it quickly damps.

To add a proportional or derivative link to the regulator is not practical, since this only increases the reverse overshoot. In particular, this reverse overshoot, depending on the coefficient of the derivative link, is shown in Fig. 14.20.

The process shown in *Fig*. 14.19 (middle line) is the most satisfactory one. Despite the fact that the external circuit has increased the duration of the transient to a value of 25 s, it eliminated the reverse overshoot and eliminated the conventional overshoot, thus we have obtained a satisfactory result except for a small reverse overshoot that can be neglected.

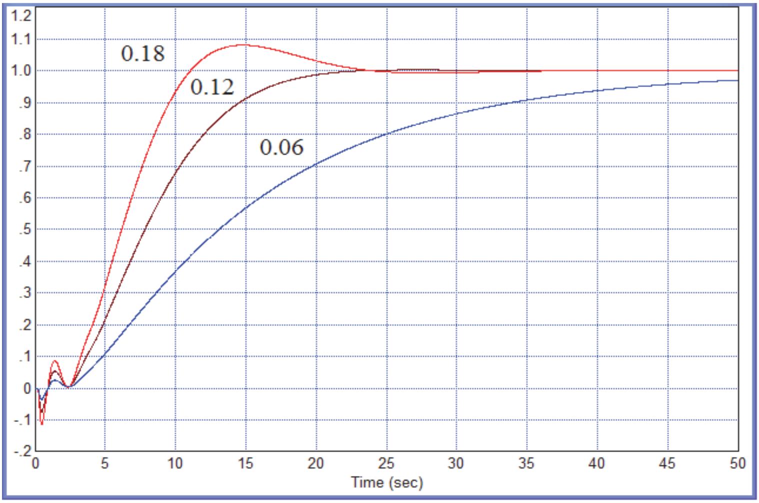


Fig. 14.19. Transient process of the system with an integrator in the external loop (the coefficient of the integrator is indicated on the graph)

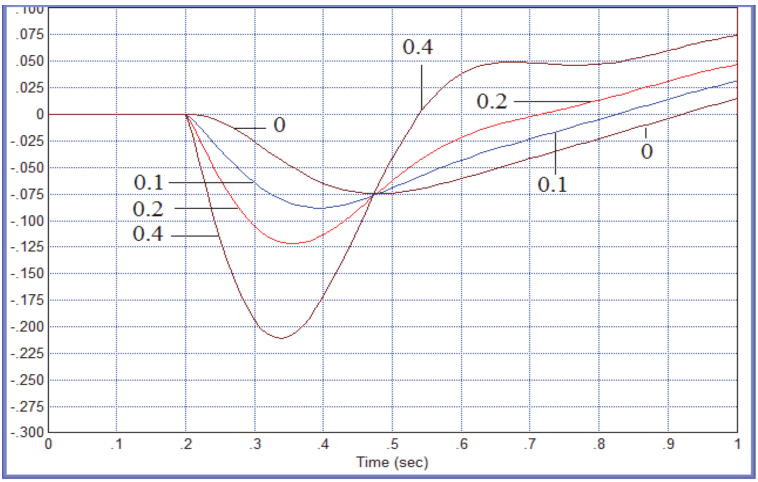


Fig. 14.20. The dependence of the initial part of the transient process on the coefficient of the derivative link (the value of the coefficient is indicated on the graph)

The proposed structure with two loops has all the properties of the structures with a single feedback and advantageously differs, for example, by the localization method, or from the structure with the reference model. In these structures, smooth control is achieved only by prescription, but by the disturbance, it does not work so effectively, whereas in the proposed system the relationship between the prescription error and the disturbance is unique: they coincide in form, as in conventional linear systems with a single feedback loop. To illustrate this fact, Fig. 14.21 shows the transient processes with a stepped single jump of the disturbance applied to the output of the object in an additive form. The graph combines two processes, starting, respectively, at instants $t = 0.1$ s and $t = 20$ s. The form of these two graphs is identical, which indicates the robustness of the system. In addition, the form of these graphs is completely identical to the type of error when applying a single stepwise influence on the input of the system (the sum of the graph in Fig. 14.21 for $t = 0.1$ s and the graph in Fig. 14.19 is equal to unit, according to one of the basic relations of the theory for single-loop systems with a unit negative feedback, as it can be easily seen).

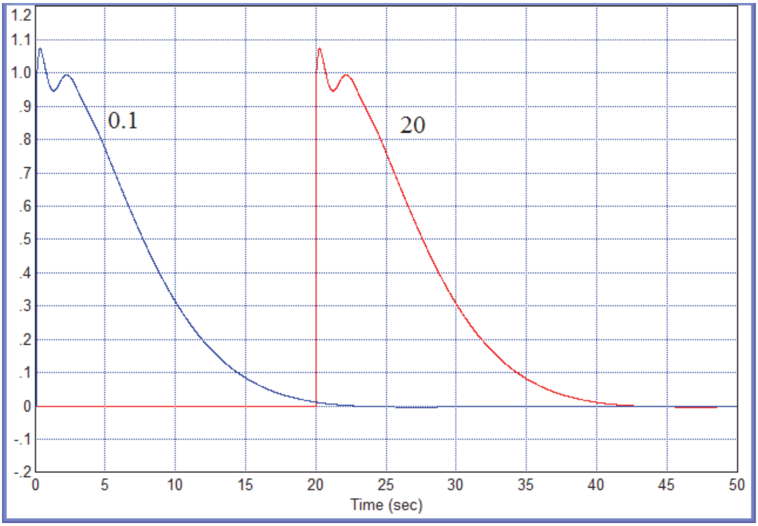


Fig. 14.21. Response of the composite system to a single jump of the disturmance with delay (the delay value is indicated on the graph)

14.4. Discussion

In this paper, known methods of numerical optimization of the regulator are applied to an extremely complex object with a mathematical model of the form (14.1). The object is prone to increasing oscillations at the output, even with the smallest impact on its input. The control of such an object is highly nontrivial. The methods of the Bypass channel, the method of pseudo-local loops using the open model, the method of additional differentiation (PIDD-regulator) are investigated.

It is shown for the first time that the value of the delay of the input signals can be critical for the accuracy of the simulation, namely: modeling with zero delay gives results very different from the results for nonzero delay. In this case, of nonzero delay, all the results are identical. On this basis, it is proven that in simulation, a non-zero delay of all test signals should be used. This is because at the beginning the simulation software must reliably simulate the start conditions, only after that the signals in the system can begin to change so that the program simulates the dynamic development of the signals in the system.

It is also shown that the use of an open model of the system does not provide reliable robust solutions, since differences in different models can be critical when steps numerically model them. Since steps can only carry out any simulation in practice, this limitation is insurmountable.

These features allow for ensuring the accuracy and robustness of modeling and optimization.

It is also shown that the method of numerical optimization, while ensuring its accuracy, even for oscillatory unstable object, gives a reliable result that provides stability, although it is characterized by a significant reverse overshoot.

A method is proposed for an additional external loop with an integrator as a sequential regulator and with unit negative feedback. It is shown that, although the numerical optimization of the coefficient of this integrator may be ineffective, the empirical adjustment of this coefficient is extremely simple (one-dimensional optimization can be performed by the method of dichotomous division/multiplication). The result is characterized by a sufficiently high quality, robustness and equally effective suppression of the error not only in regard to presctiption, but also in regard to disturbance.

14.5. Conclusions for the chapter

As a result of this research, additional important recommendations on the use of the numerical optimization method have been developed. In particular, recommendations concerning the requirement of robustness of the regulator have been made, a requirement, which is critical for practical applications. On the basis of modeling, shortcomings of the method of pseudolocal loops using an open model are revealed, since the identity of open and closed models can be violated in practice despite the fact that the mathematical description of such models is completely identical. This is due to the approximation of the simulation in steps and the fact that in the actual implementation of the regulator only such a simulation can be used.

The method of an additional external loop is proposed and tested. The effectiveness of the method was proved and demonstrated.

15. PRACTICAL RECEARCH AND LABORATORY WORKS IN THE FIELD OF FEEDBACK CONTROL IN TECHNICAL UNIVERSITIES OF LIBEREC AND SOFIA

Laboratory works in Technical University of Liberec in the field of robotics and machatronics are closely connected with the task of feedback control.

The following photos illustrates an equipment of laboratories for practical training of students, some results (original mobile robots, electromobile designed by students, control of some chemical processes, modeling of the heat distribution) and big interest of students to practical research, related to the advanced technologies in automotive industry or in power plants equipment (eg Škoda Auto or Nuclear power plant Temelín).

Photo 1. Nuclear power plant Temelín

Photo 1–13 show laboratory works in TUL. Photo 14–21 show laboratory works in TUL and constructed there robots.

Photo 2. Nuclear power plant Temelín and students of TUL

Photo 3. Electro mobile developed in TUL with the helps of students

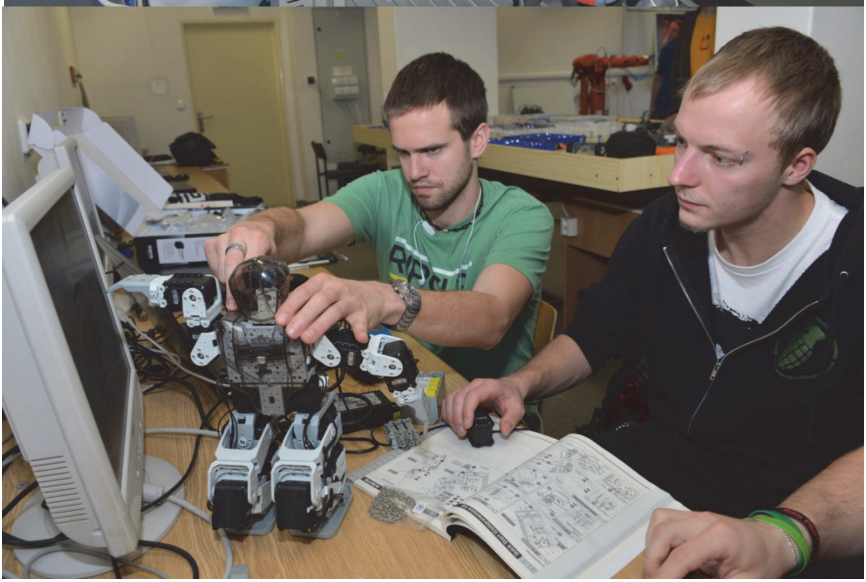
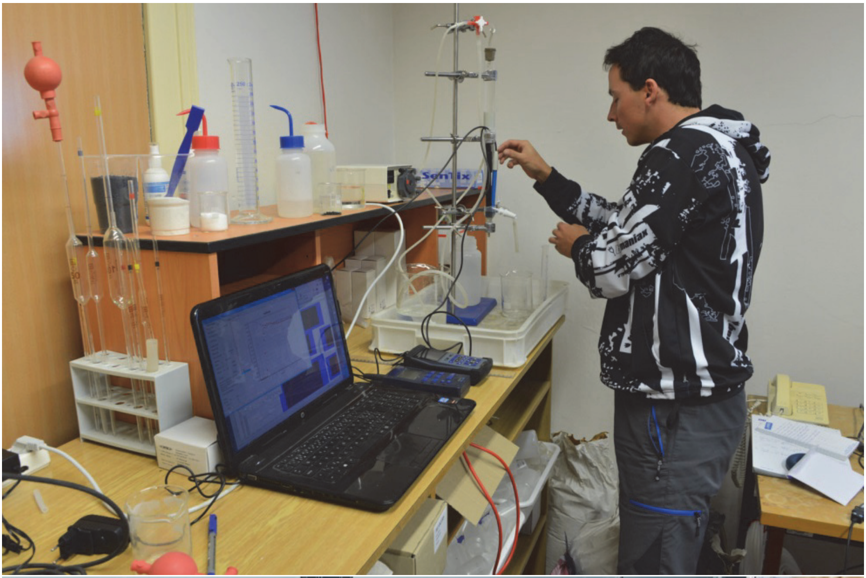


Photo 4 and 5. Students developing robotic device

Photo 6 and 7. Students developing robotic device



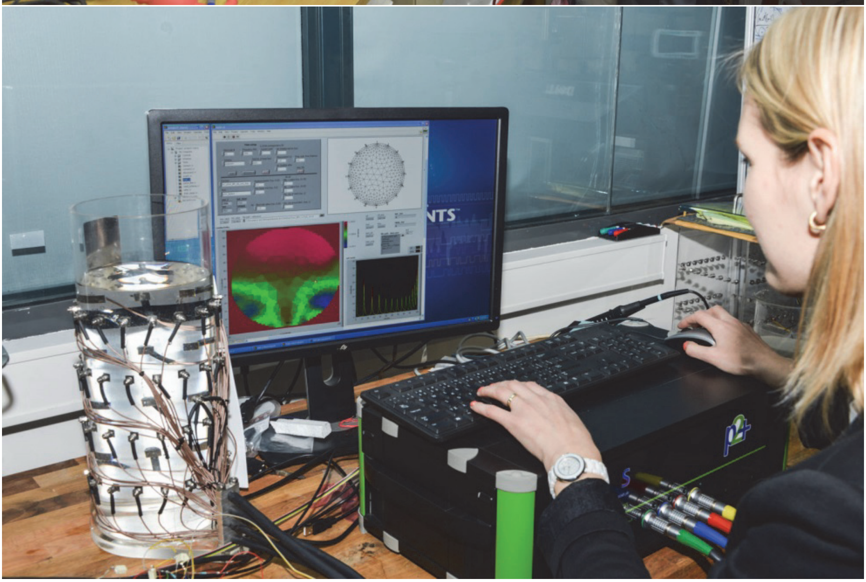
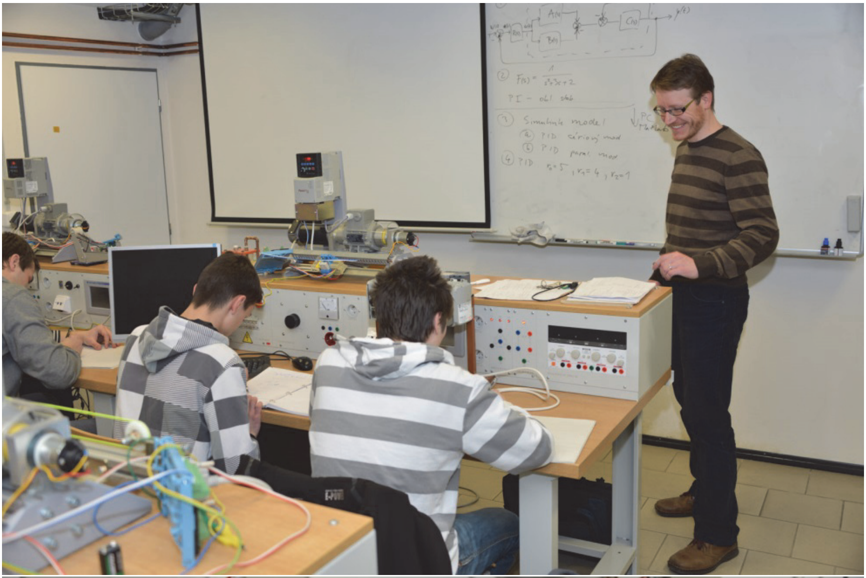


Photo 8 and 9. Students of TUL doing laboratory works



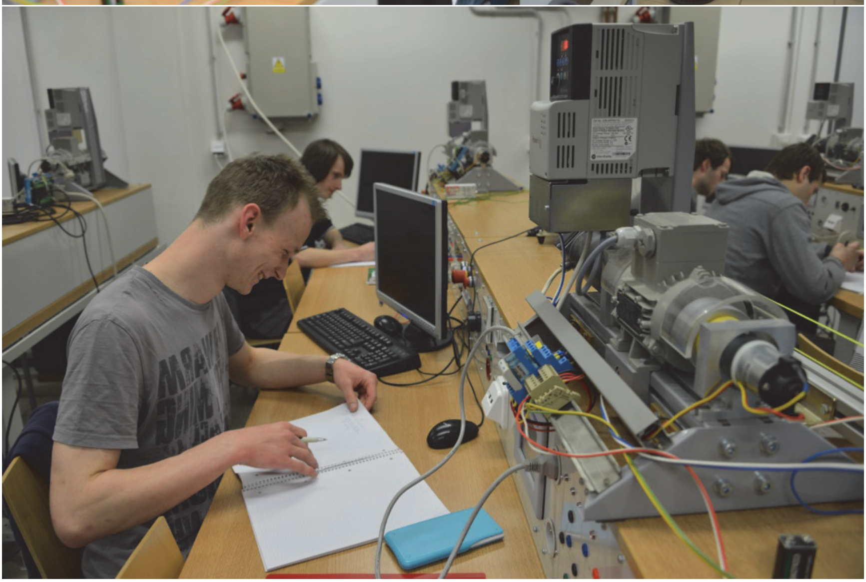


Photo 10 and 11. Students of TUL doing laboratory works

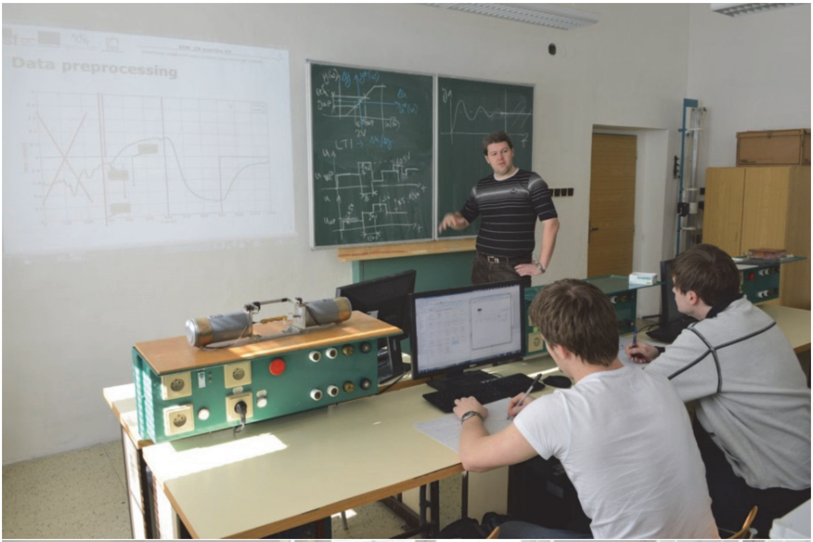
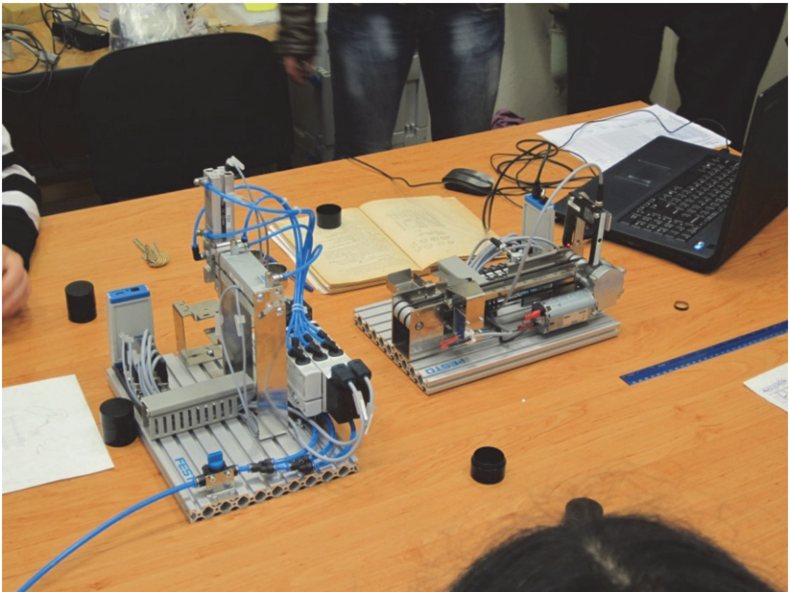


Photo 12 and 13. Students of TUL doing laboratory works



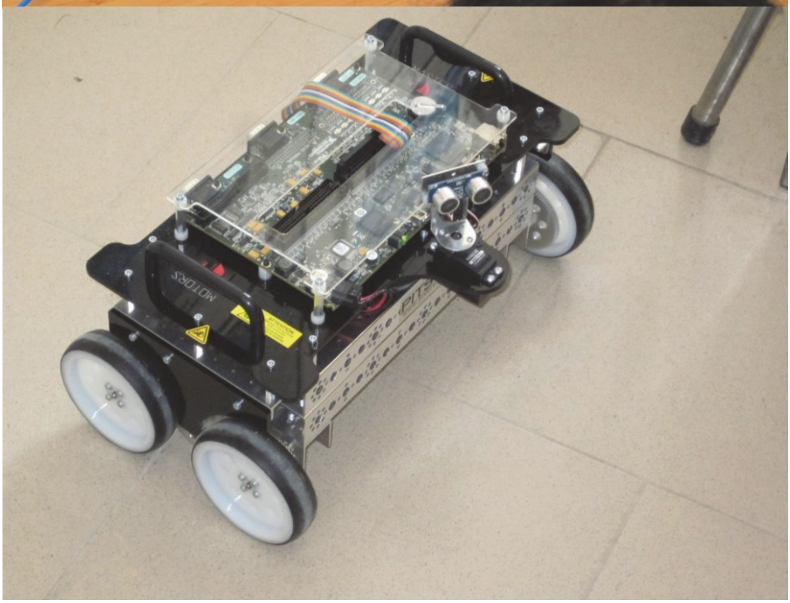
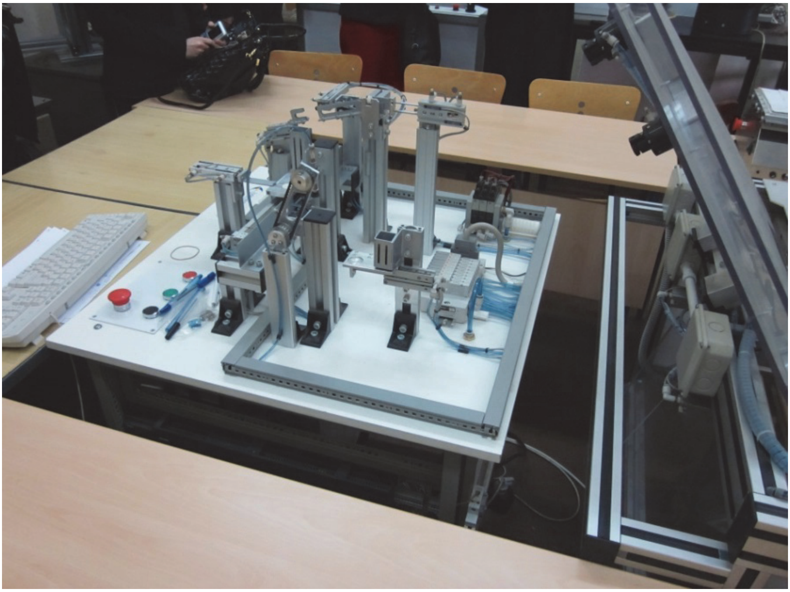


Photo 14 and 15. Robots constructed in TUS

Photo 16 and 17. Laboratory works on robotics in TUS



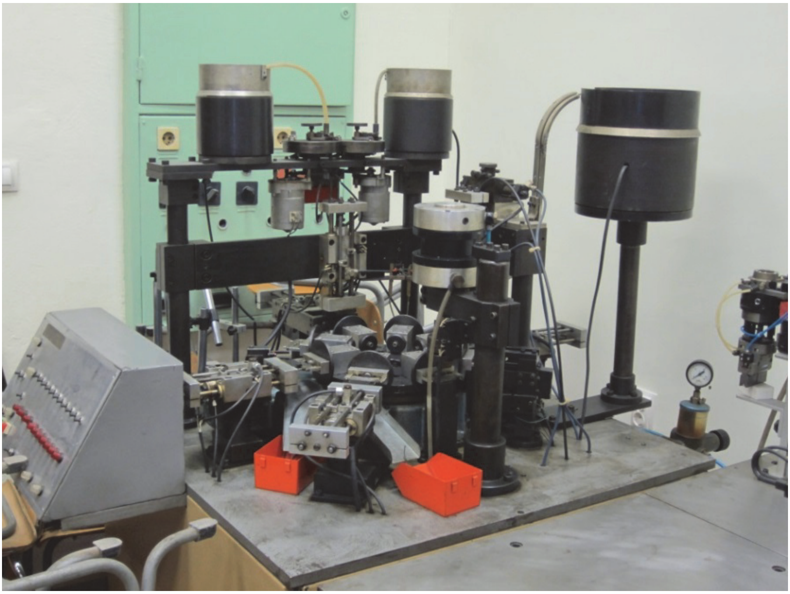


Photo 18 and 19. Robots constructed in TUS

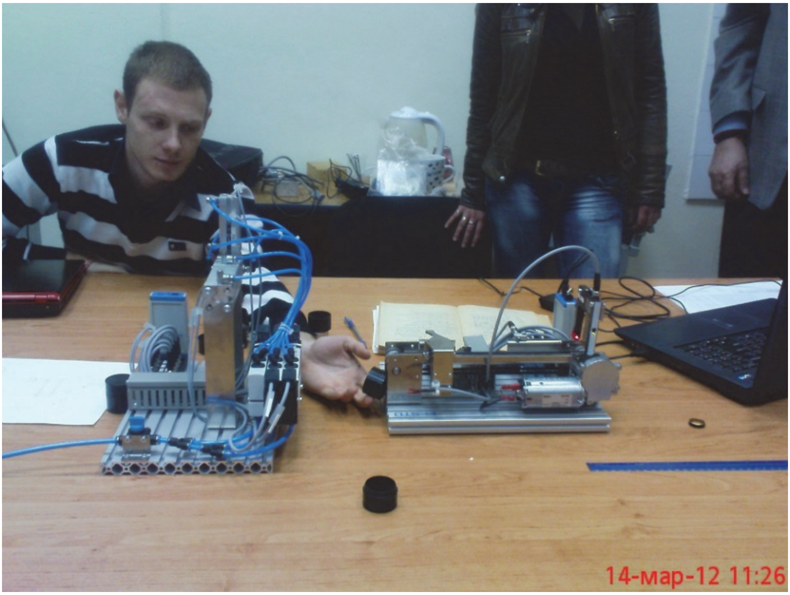


Photo 20 and 21. Laboratory works on robotics in TUS

16. QUESTIONS FOR CONTROL AT EXAMS

1. What is the difference and what is the similarity of the structures of regulators with the bypass channel and with the predictor Smith? Show this difference on the structural diagrams and on the features of the model.

2. How can we compare the effectiveness of several different structures of the regulator in order to exclude the influence of the quality of the settings of these structures? Give an example.

3. What is the purpose of using two control channels for an object with one output? Give an example.

4. What positive qualities of various sensors can be connected and what disadvantages are compensated for? Give an example with a block diagram.

5. What properties of a closed system are affected by the displacement of the sensor? What qualities of a closed system do sensor noise affect? Give examples of transient processes of the respective systems.

6. What are the subtasks of the optimization task of the regulator (it is necessary to list them, following the sequence of actions when solving them)?

7. What can hamper the iterative optimization of the regulator in practice?

8. Is it theoretically possible to optimize the PI controller settings for a first-order object? How to understand the impossibility of optimizing the PI (or PID-) controller for a specific object model?

9. What does a "robust" and "non-robust" solution mean? How can one be convinced of a non-robust solution?

10. What causes noise when the signal is repeatedly derivated in a digital way?

11. What are composite cost functions used for? Give an example.

12. How is it necessary to modify the cost function to limit the overshoot? Is there a single way for this solution?

13. What is the term "convergence of optimization algorithms"? Give an example.

14. What determines the convergence of the solution of the problem by this optimization algorithm? Give an example.

15. What is a "locked system quality criterion"? What is it for? Give an example.

16. What are the requirements for the quality criteria of closed systems?

17. What cost functions as applied to the optimization problem for regulators do you know? What are their similarities?

18. Why is it necessary for the cost function to combine several additive components (when necessary)? Give an example.

19. What additional restrictions can be implemented by introducing additional terms into the cost function? Give examples.

20. Give examples of the feasibility of introducing non-linear elements into the controller.

21. How changes speed, overshoot and static error in a system with a first-order object and a proportional controller change with increasing gain? Explain on the charts.

22. How changes speed, overshoot and static error in a system with a second-order object and a proportional controller change with increasing gain? Explain on the charts.

23. How changes speed, overshoot and static error in a system with a third-order object and a proportional controller change with increasing gain? Explain on the charts.

24. How changes speed, overshoot and static error in a system with a first-order object and a PI controller change with increasing gain? Explain on the charts.

25. How changes speed, overshoot and static error in a system with a second-order object and a PI controller change with increasing gain? Explain on the charts.

26. How changes speed, overshoot and static error in a system with a third-order object and a PI controller change with increasing gain? Explain on the charts.

27. How changes speed, overshoot and static error in a system with a first-order object and a PID controller change with increasing gain? Explain on the charts.

28. How changes speed, overshoot and static error in a system with a second-order object and a PID controller change with increasing gain? Explain on the charts.

29. How changes speed, overshoot and static error in a system with a third-order object and a PID controller change with increasing gain? Explain on the charts.

30. List what needs to be remembered when modeling systems based on nonlinear objects. Give an example of incorrect modeling.

31. From what considerations is the integration step chosen? How to make sure that the integration step is small enough? What can bring an insufficiently small integration step?

32. From what considerations is the integration method chosen? What method of integration is recommended to choose in the program VisSim? What can the result of incorrectly chosen integration method?

33. Optimization of regulators for multichannel objects: the structure of the object (with the linear part and the delay), the structure of the regulator, the objective functions and test input signals. Explain by example.

34. Optimization of regulators for multichannel objects: principles for choosing the numbering of channels, changing the numbering, the method of separation of movements according to the pace (when it is enough to use separate contours). Show it with examples (draw a block diagram for optimization).

35. Optimization of regulators for multichannel objects: a bypass channel for this case, principles (rules) of using a bypass channel. Explain on an example with a structural scheme.

36. Expediency of a full or incomplete PID controller for multichannel objects. Explain on examples with structural schemes.

37. Dependence of the optimization result on test signals (linearly increasing, stepped, and so on). Interrelation of the test signal selection with the required static system characteristics. Explain on examples. What is a complex test signal and how is it applied?

38. Features of the action of the objective functions in the design of static and astatic systems, as well as features of system optimization in terms of noise (with a real object or with a noisy model). Explain on examples.

39. Measures to eliminate the shortcomings of the solutions obtained (addition of objective functions). Explain on any two examples.

40. Design of robust regulators by the method of simultaneous synthesis of two or more systems. Explain on examples.

41. Designing piecewise-robust regulators. The technique. The advantages of this method. Explain on examples.

42. Regulator with pseudolocal connections. Explain with examples, explain its advantages and disadvantages.

43. Disclosed object model and its use for designing a regulator with pseudolocal connections. The choice of integration method for modeling the disclosed object models.

CONCLUSLION

In this book, we consider methods for calculation of regulators for controlling linear and nonlinear objects. Since analytical methods are inapplicable in the case of nonlinear objects, and even in the case of linear objects of great complexity, for example with elements of pure delay, numerical optimization methods are by now the only tools for solving these problems. The book consistently outlines the features of systems with negative feedback, gives criticism to table methods, arguments for preference of the VisSim software as compared to other software packages, such as MATLAB Simulink and Mathcad. It also formulates the basic requirements to a system and the mathematical apparatus applied and gives the requirements for the physical feasibility of the model. The selection of the regulator structure is justified. The current book also provides comments focused on regulators with fractional integration or differentiation. A classification of methods for synthesis of automatic control systems is given as well. The authors propose a new classification of adaptive and self-adjusting systems and formulates the differences between them. Also, the book considers the problems of choosing the method of integration, which is extremely important when implementing digital regulators in practice. For the first time the authors pay attention to the fact that the peculiarities of the operation of digital regulators depend on the method of calculating integrals and derivatives of functions obtained in digital readings, and this is correlated with the method of integration in numerical simulation and optimization. The authors analyze the reasons for the difference in the results of theoretical analysis, modeling, and practice. These analyses help to avoid such differences, which ensure that the basic properties of transient processes in a system coincide with the simulation results and correspond to the real system. The requirements to the cost function and the methods of forming such functions from several terms are considered as well.

This book considers a method for dividing motions theoretically and confirms it by modeling, including two of its modifications: the application of several drives and the combination of the advantages of different sensors, as well as the combination of these principles in one structure. A method of energy saving is proposed. The book suggests a method based on the use of the error growth detector and a method for using the bypass channel, which is the development of the idea of the Smith predictor [45–48].

The authors describe a method for providing robust control and a method for developing piecewise adaptive control systems.

They suggest methods for numerical optimization of multichannel regulators for controlling multi-channel objects. They suggest the use of the bypass channel method to this problem. It is shown that the error growth detector makes possible to provide better transient processes in the case of control of multi-channel objects.

The authors solve the problems of control of objects prone to oscillations. They also solve the task for controlling of multichannel object with relatively high dimensionality, particularly 3 × 3. The considered methods are extremely effective for many control problems such as those described in [23–43].

In this book, we do not describe the control technique using local and pseudo-local feedbacks. We do not give detailed classification of adaptive systems and authorial structures of adaptive systems and do not consider systems with competitive quality criteria. These problems are discussed elsewhere (for instance, in papers published in the journal "Automatics & Software Engineeery"); see [23–43].

Acknowledgments

The authors extend their gratitude to all their colleagues, teachers, and students who contributed to this work with support and advice. A special word of thanks goes to Professors S. N. Bagaev, A. S. Vostrikov, G. A. Frantsuzova, A. A. Voevoda, Associate Professors G. V. Sablina, A. B. Kolker, V. M. Semibalamut, A.V. Liapidevskiy, and some others.

REFERENCES

1. Astrom K. J., Hagglund T. Advanced PID-control. ISA (The Instrumentation, Systems and Automation Society), 2006. 460 p.

2. Doyle J.C., Glover K., Khargonekar P.P., Francis B.A. State-space solutions to standard H2 and H∞ control problems // IEEE Transactions on Automatic Control AC-34, № 8, 1989.

3. Grimble M. J. GMV control of non-linear continuous-time systems including common delays and state-space models. Int. Journ. of Control. 2007. P. 150–165.

4. Havre K., Skogestad S. Achievable performance of multivariable systems with unstable zeros and poles // International Journal of Control. 2001. V. 74. N. 11. P. 1131–1139.

5. Ilchmann, Achim and Ryan, Eugene P. Asymptotic tracking with prescribed transient behaviour for linear systems. Int. Journ. of Control. 2006 P. 910 – 917.

6. Kariwala V., Skogestad S., Forbes J. F. and Meadows E. S. (2005) Achievable input performance of linear systems under feedback control. Int. Journ. of Control. P. 1327–1341.

7. Kharitonov V. L. Lyapunov matrices for a class of neutral type time delay systems. Int. Journ. of Control. 2008. P. 1–17.

8. Okou F. A., Akhrif O. and Dessaint L. -A. Decentralized multivariable voltage and speed regulator for large-scale power systems with guarantee of stability and transient performance. Int. Journ. of Control. (2005) P. 1343–1358.

9. Pan Y. -J., Marquez H. J. and Chen T. Stabilization of remote control systems with unknown time varying delays by LMI techniques. Int. Journ. of Control. (2006) 79:7, 752–763.

10. Pepe P., Jiang Z.-Po and Fridman E. A new Lyapunov-Krasovskii methodology for coupled delay differential and difference equations. Int. Journ. of Control. 2007. P. 107–15.

11. Pommaret J.-F. and Quadrat A. (2004) A differential operator approach to multidimensional optimal control. Int. Journ. of Control. P.821 – 836.

12. Rossiter J. A., Kouvaritakis B. and Cannon M. (2005) An algorithm for reducing complexity in parametric predictive control. Int. Journ. of Control. P. 1511–1520.

13. Sofrony J., Turner M. C. and Postlethwaite I. Anti-windup synthesis using Riccati equations. Int. Journ. of Control. 2007. P.112 – 128.

14. Towaim T., Barton A. D., Lewin P. L., Rogers E. and Owens D. H. (2004) Iterative learning control - 2D control systems from theory to application Int. Journ. of Control. P. 877–893.

15. Zhai G., Kondo H., Imae J. and Kobayashi T. Hybrid static output feedback stabilization of two-dimensional L TI systems: a geometric method. Int. Journ. of Control. 2006. P.982 – 990.

16. Zhang L., Shi P. and Boukas E.-K. H∞, output-feedback control for switched linear discrete-time systems with time-varying delays. Int. Journ. of Control. 2007. P. 1354–1365.

17. Bendame M. The MathCAD Advisor: Using MathCAD to Model a PID Controller http://www.mathcad.com/support/item/02aug_id.asp

18. Havre K., Skogestad S. Achievable performance of multivariable systems with unstable zeros and poles. International Journal of Control. 2001. Vol. 74. N. 11. P. 1131–1139.

19. URL: http://model.exponenta.ru/tau_lec.html.

20. URL: http://www.vissim.nm.ru/download.html .

21. V.A. Zhmud, S.V. Barmasov, V.D. Gitel'son. An electronic system for stabilizing of the frequency of the He-Ne laser to the methane absorption lines. Instruments and Experimental Techniques. 1999. T. 42. № 4. P. 551–557.

22. A phase-locked loop system for the difference frequency of two lasers. Barmasov S.V., Zhmud' V.A. Instruments and Experimental Techniques. 2000. T. 43. № 3. P. 381-383.

23. The modeling tests of the new PID-regulators structures. Voevoda A.A., Zhmud V.A., Ishimtsev R.Y., Semibalamut V.M. 2009. Proceedings of the IASTED International Conference on Applied Simulation and Modelling, ASM 2009. P. 165 – 168.

24. V. Zhmud, A. Liapidevskiy, E. Prokhorenko. The design of the feedback systems by means of the modeling and optimization in the program VisSim 5.0/6. 2010. Proceedings of the IASTED International Conference on Modelling, Identification and Control. P. 27–32.

25. Zhmud V.A., Liapidevskiy A.V. The Design of the Feedback Systems by Means of the Modeling and Optimization in the Program VisSim 5.6/6.0 // Proc. Of The 30th IASTED Conference on Modelling, Identification, and Control ~ AsiaMIC 2010 ~November 24–26, 2010 Phuket, Thailand. PP. 27–32.

26. V. A. Zhmud. The Use of the Feedback Control Systems in Laser Physics Researching Experiments. // Proceedings of RFBR and DST Sponsored “The 2-nd Russian-Indian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics”, 10 – 13 September, 2011, Additional volume, pp. 40–43.

27. V. Zhmud, A. Polishchuk, A. Voevoda, R. V. Rao. The Tuning of the PID-Regulator for Automatic Control System of Thermo Energetic Equipment // Proceedings of the Fifth International Conference on Advances in Mechanical Engineering (ICAME-2011), June 06-08, 2011. Surat – 395 007, Gujarat, India. pp. 254–263.

28. V.A. Zhmud. The Use of the Feedback Control Systems in Laser Physics Researching Experiments. Proceedings of RFBR and DST Sponsored "The 2-nd Russian-Indian Joint Workshop on Computational Intelligence and Modern Heuristics in Automation and Robotics", 10–13 September, 2011, Additional volume. P. 40–43.

29. Modern key techologies in automatics: Structures and numerical optimization of regulators. Zhmud, V., Yadrishnikov, O., Poloshchuk, A., Zavorin, A. 2012. Proceedings - 2012 7th International Forum on Strategic Technology, IFOST 2012.

30. V. Zhmud, O. Yadrishnikov, A. Poloshchuk, A. Zavorin, Modern key techologies in automatics: Structures and numerical optimization of regulators. 2012. Proceedings – 2012 7-th International Forum on Strategic Technology, IFOST 2012.

31. V. Zhmud, O. Yadrishnikov. Numerical optimization of PID-regulators using the improper moving detector in cost function. Proceedings of the 8-th International Forum on Strategic Technology 2013 (IFOST-2013), vol. II, 28 June – 1 July. Mongolian University of Science and Technology, Ulaanbaator, Mongolia. IEEE organized. 2013. P. 265 – 270. http://www.must.edu.mn/IFOST2013/

32. Zhmud V.A., Zavorin A.N. Metodi di ottimizzazione del controllo numerico su una modelli troncati. Italian Science Review. 2014; 4(13). PP. 686-689. Available at URL: http://www.ias-journal.org/archive/2014/april/Zhmud.pdf

33. V.A. Zhmud, A.N. Zavorin. Method of designing energy-efficient controllers for complex objects with partially unknown model. Proc.: The control and modeling in complex systems. Proceedings of the XVI International Conference June 30–July 3, 2014, Samara. Russia. P. 557–567.

34. V. Zhmud, L. Dimitrov. Designing of complete multi-channel PD-regulators by numerical optimization with simulation. 2015, 2015 International Siberian Conference on Control and Communications, SIBCON 2015 – Proceedings.

35. V.A. Zhmud, O.D. Yadrishnikov, V.M. Semibalamut. Control of the objects with a single output and with two or more input channels of influence. WIT Transaction on Modelling and Simulation. Vol.59, 2015. WIT Press. www.witpress.com. P. 147–156.

36. V.A. Zhmud V., L.V. Dimitrov. The providing of the power saving control of one output value with two controlling channels having different effectiveness and cost of the controlling resource. International journal for science, technics and innovations for the industry. 2015. IX Issue 9. ISSN 1313-0226 / Machines, Technologies, Materials. P. 50 – 54.URL: http://mech-ing.com/journal/Archive/2015/9-2015.pdf

37. Vostrikov A.S., Frantsuzova G.A. Synthesis of PID controllers for nonlinear nonstationary plants // Optoelectronics, Instrumentation and Data Processing. – 2015. – Vol. 51, iss. 5. – P. 471–477.

38. Zemtsov N. S., Frantsuzova G.A. Design of PID-Controller for a Once-Through Boile //Mekhatronika, Avtomatizatsiya, Upravlenie, 2015, Vol. 16, no. 9, pp. 631–636. DOI: 10.17587/mau.16.631-636

39. A. Ivoilov, V. Trubin, V. Zhmud, L. Dimitrov. Vadim Detection of Unrevealed Non-linearities in the Layout of the Balancing Robot. Proc. of International Siberian Conference on Control and Communications (SIBCON). Moscow, RUSSIA. MAY 12-14, 2016. IEEE International Siberian Conference on Control and Communications.

40. Zhmud, Vadim; Dimitrov, Lubomir. Investigation of the Causes of Noise in the Result of Multiple Digital Derivations of Signals Researches with Mathematical Modeling. 11th International IEEE Scientific and Technical Conference on Dynamics of Systems, Mechanisms and Machines (Dynamics). Omsk, RUSSIA: NOV 14-16, 2017.

41. Zhmud, Vadim; Hardt, Wolfram. Practical Dependance of the Error of ADC on Conversion Frequency Researches with Mathematical Modeling: 11th International IEEE Scientific and Technical Conference on Dynamics of Systems, Mechanisms and Machines (Dynamics): Omsk, RUSSIA: NOV 14-16, 2017.

42. Zhmud, V. A.; Dimitrov, L. V.; Roth, H. A New Approach to Numerical Optimization of a Controller for Feedback System. Destech Publicat Inc.: 2nd International Conference on Applied Mechanics, Electronics and Mechatronics Engineering (AMEME): Beijing, PEOPLES R CHINA.: OCT 22-23, 2017

43. Zhmud V.Z., Zavorin A.N. Compensation of the Sources of Unwanted Direction of the Transient Process in the Control of Oscillatory Object. Automatics & Software Engineery. 2013. N3. http://jurnal.nips.ru/sites/default/files/ASE-3-2013-3.pdf

44. V. Zhmud. Systems of automatic control. Precision control of laser light. (In Russian) Moscow. Urait. 2018. http://urait.ru/catalog/412103

45. Smith predictor. URL: https://en.wikipedia.org/wiki/Smith_predictor

46. Control of Processes with Long Dead Time: The Smith Predictor. URL: https://www.mathworks.com/help/control/examples/control-of-processes-with-long-dead-time-the-smith-predictor.html

47. Overcoming process deadtime with a Smith Predictor. URL: https://www.controleng.com/single-article/overcoming-process-deadtime-with-a-smith-predictor.html

48. Control Systems, Robotics, and Automation. Vol. II. Smith Predictor and Its Modifications. Hang C. C. www.eolss.net/sample-chapters/c18/e6-43-03-05.pdf

49. Dyakonov V. VisSim+Mathcal+MATLAB. Visual mathematic modeling. Moscow. Solon-Press. 2004. (In Russian. Original name: Дьяконов В. VisSim+ Mathcal+MATLAB. Визуальное математическое моделирование. Москва. Солон-Пресс. 2004). ISBN 5-98003-130-8.

APENDIX

The literature recommended by the Technical University of Liberec for subjects of the study fields "Automatic Control and Engineering Informatics" and "Mechatronics".

MTI/CRI Digital Control

- Franklin, G., F. & Powell, J., D. & Workman, M., L. *Digital Control of Dynamic Systems.* Addison - Wesley, 1998.
- Albertos, P., Sala, A. *Multivariable Control Systems*. Springer Verlag,, 2004. ISBN 1852337389.
- Ljung, L.: *System Identification. Theory for the User. Prentice - Hall, Inc. Upper Saddle River*. New Jersey, 1999.
- Rossiter, J.A. *(2003), Model-based predictive control, A Practical Approach, CRC Press*.

MTI/RTC Real Time Control Systems

- Wittenmark, B., Aström, K. J., Arzén K. E. *Computer Control: An Overview*. IFAC Professional Briefs, 2007.
- Landau, Ioan D., Zito, Gianluca. *Digital Control Systems Design, Identification and Implementation*. Springer Verlag, 2006.
- Liu, J. W. *Real Time Systems*. Prentice Hall, Upper Saddle River, N.J., 2000.

MTI/STR State Space Control Systems

- Luenberger, D.G. An Introduction to Observers. IEEE Trans. on Automatic Control, Vol. AC-16, pp.596-602.
- Raven, F.H. Automatic Control Engineering. NY, McGraw-Hill, Inc. 1995.
- Kuo, Benjamin C. Automatic Control Systems. London, Prentince-Hall, Inc. 1991.
- Van de Vegte, J. Feedback Control Systems. NJ, Prentice-Hall, Inc. 1990.
- Kwakernaak - Sivan: Linear Optimal Control Systems. NY, John Wiley & Sons, Inc. 1972.

MTI/ARI Automatic Control

- Kuo, B.C.: *Automatic control systems.* John Wiley & Sons, Inc. New York, 1995.
- Grace, A., Laub, J., A., Litle, J., N., Thomson, C., M. *Control System Toolbox. For Use with MATLAB. User's Guide.* The Math Works,Inc., 1995.

- Grace, A.-Laub,J.A.-Little,J.N.-Thompson,C.M. *Control System Toolbox. For Use with MATLAB. User's Guide.* The Math Works,Inc., 1995.
- Isermann, R.:. *Mechatronics Systems. Fundamentals..* Springer, London., 2003. ISBN ISBN 1852336935.
- Shinskey, F. G. *Process Control Systems. Application, Design, and Tuning*. Mc Graw Hill. ISBN 0-07-057101-5.

MTI/RVS Multivariable Control Systems

- Goodwin, G., C., Graebe, S., F., Salgado, M., S. *Control System Design*. Prentice Hall, 2001. ISBN 139586539.
- Phillips, Ch., L., Harbor, R., D. *Feedback Control Systems*. Prentice Hall, 2000. ISBN 133736970.
- Albertos, P., Sala, A. *Multivariable Control Systems*. Springer Verlag,, 2004. ISBN 1852337389.
- Skogestad, S., & Poslethwaite, I.: *Multivariable Feedback Control. Analysis and Design.* John Wiley & Sons, 1996. ISBN 470011688.
- Marlin, T., E. *Process control. Designing processes and control systems for dynamic performance.* McGraw Hill, 2000. ISBN 70404917.

MTI/OPT Optimization Problems

- R.Fletcher. *Practical methods of optimization.* 1987. ISBN ISBN 0-471-91547-.

MTI/IDS System Identification

- Schoukens, J., Pintelon, R., Rolain, Y. *Mastering System Identification in 100 Exercises*. IEEE Press-Wiley, 2012.
- Van Overschee, P., De Moor, B. *Subspace Identification for Linear systems: Theory - Implementation - Applications*. Kluwer Academic Publishers, Leuven, 1996.
- Soderstrom, T., Stoica, P. *System Identification*. Prentice Hall, Uppsala, 2001. ISBN 0-13-881236-5.
- Schoukens, J. *System Identification*. Brusel, 2013.

MTI/ASR Algebraic Method Design

- Mordell, L.J. Diophantine Equations. Academic Press, London 1969.
- Chen, C.T. Linear Systems Theory and Design. Holt, Rinehart and Winston, New York 1984.
- Barnett, S. Polynomial and Linear Control Systems. Pure and Applied Mathematics, Marcel Dekker, New York 1983.
- Polynomial Toolbox Manual. PolyX,Ltd., Prague 1999.

MTI/RSM Control Systems in Mechatronics

- Kuo, B.C.: *Automatic control systems.* John Wiley & Sons, Inc. New York, 1995.
- Grace, A.-Laub,J.A.-Little,J.N.-Thompson,C.M. *Control System Toolbox. For Use with MATLAB. User's Guide.* The Math Works, Inc., 1995.
- Isermann, R.: *Mechatronics Systems. Fundamentals.* Springer, London., 2003. ISBN 1852336935.
- Shinskey, F. G. *Process Control Systems. Application, Design, and Tuning*. Mc Graw Hill. ISBN 0-07-057101-5.

Vadim A. Zhmud, Lubomir V. Dimitrov, Jaroslav Nosek

NUMERICAL OPTIMIZATION OF REGULATORS FOR AUTOMATIC CONTROL SYSTEM

Textbook

In the authors' Edition

Signed Print 25.02.2019
Format 60 × 84 1/16. Newsprint
Uch.-ed. 1. 17,2. Pecs. 1. 18,5
Quantity 200 copies. Order No. 387

Publisher of the Novosibirsk State Technical University
630073, Novosibirsk, Russia, 20, Karl Marx prospect
Tel: (383) 346-31-87
E-mail: office@publish.nstu.ru

Printed in the Printing House
of the Novosibirsk State Technical University
630073, Novosibirsk, Russia, 20, Karl Marx prospect