Proceedings of VII International Conference "System Identification and Control Problems", Moscow, 2008

Dynamic principles of prognosis and control

G.A.Leonov

Saint-Petersburg state university

1. Introduction

Consider dynamical systems, generated by differential equations

$$\frac{dx}{dt} = f(x) \tag{1}$$

or the difference ones

$$x(t+1) - x(t) = f(x(t)), \quad x \in R^n. \tag{2}$$

For these equations for any constant $S > 0$ we have the following obvious property of solutions

$$x(t+S, S, x_0) = x(t, 0, x_0). \tag{3}$$

This property also holds true in the case of more general descriptions of dynamical systems (from the point of view of their phase spaces or nonlinear operators acting in these spaces).

Relation (3) implies that the segment of trajectory $x(t, x_0)$, outgoing from the point x_0 at time $t = 0$, coincides with the segment of trajectory, outgoing from the point x_0 at time $t = \tau$,.

It follows that under the same conditions the physical experimental data are repeated and, therefore, we can theoretically forecast some processes and control them.

However by reason of arising the instabilities (which are an object of intensive consideration in recent years [1]), the coincidence, mentioned above, is often possible in sufficiently small time intervals $(0, \tau)$ and $(s, s+\tau)$ only.

We describe certain approaches to the prognosis and control, which are due to the general laws of instabilities arising in dynamical systems. These approaches, developed within the framework of "experimental mathematics", are based on that we do not try to construct, to identify, and to analyze the approximate models of rather complicated real dynamical objects but collect a certain experimental material connected with real models and then make use of it for the prognosis and the construction of control. The occurrence of instabilities obeys certain general regularities, the account of which results in certain general principles of the qualitative control theory. They are also discussed in the present report.

1. Informative example. How to make a weatherforecast for a week. The reason why it cannot be made for more than two weeks.

In 50-60s of the 19th century the progress in the fields of continuum mechanics and computational mathematics made it possible to propose more accurate mathematical models of atmospheric changes, to construct more effective algorithms for the solution of differential equations of these models, and to realize these algorithms with the help of more high-speed computers. By this breakthrough it occurred a widespread opinion that, having made some additional efforts in these directions, we can make the weatherforecast for many weeks, months, and even years.

However it turned out afterwards that a long weatherforecast is impossible in principle.

This fact was established theoretically in the works of E.Lorenz and his progeny [1], discovered the instability in the mathematical models of atmosphere. The latter means a strong sensitivity of solutions of differential equations, describing atmospheric processes, with respect to the initial data. The understanding of this fact gave rise to observational experiments, in the framework of new direction in "experimental mathematics", what will be considered below.

By now in Europe the large material of meteorological observations has been accumulated. Such observations have been carried out regularly for many decades.

Consider the observations, for example, in May 9, 2004 in a certain region of Europe. We choose then a certain year (for example, 18: such that in May 9, 18: in this region, approximately the same meteorological parameters (a temperature, pressure, air moisture, a strength and direction of wind, cloudiness and so on) were observed. The above parameters are used as initial (and boundary) conditions for the solutions of differential equations of atmospheric model.

These equations describe the laws of continuum mechanics, which hold always true, in any year. Therefore these equations hold also true and are the same in different years. The solutions of these equations are uniquely determined by initial data. Since the equations and the initial data in May 9, 2004 and in May 9, 18: are identical, we believe that the solutions, describing the change of meteorological data (such as a temperature, pressure, air moisture and so on), are also identical.

Hence for each day from the chosen time interval, the parameters, which had been observed for a month (from May 9, 2004 to June 9, 2004 and from Nay 9, 18: to June 9, 18:), must coincide with adequate accuracy. It would seem, therefore, that the weather conditions in June 1, 2004, for example, and in June 1, 18: must be very close to each other. However the experiments show that such a coincidence is possible only on the time intervals not exceeding two weeks. At the same time for a week the coincidence of weather conditions can be rather close and in the meteorology this fact is most often used for a short-range forecast. But in the time intervals exceeding two weeks the results of observations are highly diverged. Therefore as a rule the weather conditions in June 1, 2004 and in June 1, 18: are different.

What's the matter? It turns out that a small divergence of initial data at the initial moments of observations results in the great divergence of observable parameters already after two weeks.

Thus, *even though the mathematical model of atmosphere is sufficiently correct, the computer technique is advanced, and the computer is high-speed the only result we obtain is that the correct weatherforecast for two weeks is impossible.* For this reason Japanese refused to make a weatherforecast for more than ten days.

Here it should be remarked the following. While in conventional approach as the basic problems we regard the construction of more accuracy mathematical models, the development and realization of numerical algorithms for the solution of differential equations and for identification of parameters of these equations, in the considered approach the main problem is a generation of special databases.

2. The prognosis of market behavior as the analog of a weatherforecast.

What analogy is it between the changes of weather and market? The answer to this question is as follows.

While the physical laws and the corresponding equations of convection are valid for any time interval, the similar laws of market depend on a policy, financial circumstances, and a drive of market participants. All of them are the same only on the short time interval $[t_0, T]$ (days or hours). Clearly, the changes of the variable quantities of market $x_j(t)$ obey the laws of market. The number of such quantities (by using a comparison with a weatherforecost) must be about ten ($j = 1, \dots, 10$).

In the case of a market the varying of initial (and boundary) conditions like those, used for the weatherforecast, occurs by virtue of the change of variables on the certain "initial"time interval $[t_0, t_1]/$ (t_1 is much less than T.) This allows us to draw a certain analogue to an “initial function"for differential equations with delay.

Hypothesis. There exist classes of markets such that a "good"coincidence of all observable variable characteristics of market $x_j(t)\,(j = 1, \cdots, N)$ on the intervals $[t_0, t_1]$ and $[t_0 + \tau, t_1 + \tau](t_1 + \tau < T)$ implies their "good"coincidence on the certain intervals $[t_1, t_1 + \varepsilon]$ and $[t_1 + \tau . t_1 + \tau + \varepsilon]$, in which $t_1 + \tau + \varepsilon < T$.

Here the "good"means a certain preliminary "smoothing or averaging"of the quantities $x_j(t)$, what is similar, for example, to how we account for a certain "average"velocity of wind, by smoothing the counterblasts or weakening of it on small time intervals.

Thus, in analogy with the forecastweather we can, evidently, forecast the behavior of certain markets on small time intervals, making use of similar parameters (characteristic variables) of the previous observations.

Certainly, this hypothesis needs to be checked on concrete multiparametric markets. Besides, in this case it is necessary to choose happily (from the experiments, as before) the time scales (i.e. $t_1, T, \tau, \varepsilon$).

3. The Klausewitz principle

This principle is studied in any Academy of General Staff of any country, which looks to its safety. It is important for us that this principle permits a wide extension and can be used not only in wartime.

Now we formulate the Klausewitz principle for military operations.

Any military operation must be designed as the operation bounded in space and time. The next operation is designed with provision for the resulting information about the previous operation.

So, any war, no matter what aims it sets, must be partitioned into separate operations such that they have their own tactical aims and follow in sequence.

The aim of operation and the forces and resources, which are used for its accomplishment, are corrected with due regard the resulting information on the previous operations.

The choice of spatial and time restrictions of operation is an object of military art. They are often chosen as a result of a heavy and murderous preceding experience.

Note that in the case now being considered there is the analogy with the well-known fact that the weatherforecast more than for two weeks is impossible. Like the weatherforecast, in the complicated armed struggle it can also occur the instabilities and, therefore, the previously confirmed plans and models themselves can be the cause of destruction of military operation.

A shining example of the Klausewitz principle application is famous "Stalin's 10 drives"in 1944. Using the previous war experience, the General Staff of Red Army arrived at a conclusion that the optimal time period for each operation was 1–2 months and its spatial framework was 200-300 kilometers. For these operations the corresponding organizations, named fronts (the Leningrad, Karelian, Byelorussian, Ukrainian, and Baltic ones), were formed. As a result of combination of war experience and the correct use of the Klausewitz principle all ten operations, followed in sequence, were brightly ended: it was run the blockade of Leningrad and was set free Crimea, the south-west Ukraina (the Korcun'-Shevchenko operation), Byelorussia (the "Bagration"operation), Moldavia (the Yassko-Kishinev operation), and other regions of USSR.

At the same time the tendency: "to develop success"after the end of designed operation, i.e. to continue mechanically a motion, is often ended in disaster.

A shining example is a cruel rout of Red Army in 1920 about Warsaw and a collapse of German Army about Stalingrad.

These examples bring out clearly that the actions, which seem to be unnatural for arm laymen, such as the stopping of German Army about Dunkirk in 1940 (at that date English Army got an offing to evacuate himself over the Channel) and of Red Army about Warsaw in 1944 (when it was started Warsaw's armed insurrection) were made just in accordance with the Klausewitz principle. In both cases the previous operations were finished and it was necessary to design the new operations for Dunkerk and Varshaw to be taken and to make the corresponding preparations.

The Klausewitz principle must be taken into account also in the case of the putting in force of any global reforms (in a country, company, and public structure). In these cases it is necessary first to divide neatly the designed transformations into separate parts, to enforce, in series, separate transformations, to obtain results, and to design with their help the next transformations. Then it is necessary to obtain the results of second stage transformations and to account for them in planning the third stage of reforms. Only after such careful design, the third stage can start and so on.

What devastating contrast between the reforms process control, following from the Klausewitz principle, and the undigested actions of the administrative authority of SSSR and Russia in 80-90s of the last century!

Certainly, the choice of a depth and time of each stage of reforms is an object of administrative and economical art (as the similar parameters for the military operations are an object of military art). In addition it is necessary that the aims, which are set at each stage, were attainable (i.e. the tools and resources were sufficient for accomplishment of the posed aim).

Moreover it is important for us that, keeping the above-mentioned scheme, we does not lose sight of the global final cause. (Similar to the fact that a victory is a final global aim of military operations).

4. The Klausewitz principle in the problems with bounded resources

One of modifications of the Klausewitz principle is a segregation and sequential solution of priority problems in the case of bounded resources. Recall that the classical Klausewitz principle is a solution of basic problem via sequential solutions of specially segregated subproblems.

Below we shall give one of the most shining examples of applying the modified Klausewitz principle.

In the result of World War I, Germany suffered a cruel defeat. Its main co-belligerent, Austria-Hungary, was partitioned and, therefore, already was not a powerful European State. On Germany it was imposed the enormous contributions and hard restrictions on creating the modern army. In the twenties of the last century Germany was a poor and weak State. But all at once, in 1934-1936 years, it created a powerful Army equipped by advanced armament. In 1936 the army of Germany has exceeded in its power the arms of England and France together.

How came it?

After the defeat in World War I, Germany has kept the kernel of General Staff, consisting of the well-educated officers with a militant experience and a wide experience of organizing and mobilizing work. They helped to save the old design collectives and to create the new ones.

The design collectives developed and produced the test samples of new military technique. The test samples passed the tests only. The commercialization was omitted. Then, using the obtained results and recommendations of General Staff, the designers produced, at once, the next generation of arms.

This made it possible to develop secretly the best models of arms and, practically at a time in 1934, to commercialize these models, to mobilize the army, and to munition it by this arms within very short time. After that the recollected west allies arrived at a conclusion that the war with Germany in 1937 was evidently a forlorn hope.

This example shows once more the necessity of applying the the principles of control, mentioned above.

5. The "master-slave"principle

The “master-slave"principle is usual in a modern technique. In this case there is a set of similar devices ("slaves"), which are not related together and operate, at a time, complying with signals of standard device ("master") only.

For example, switching a television, you switch together the "sitting"in your television "slave", a horizontal generator, which governs a motion of beam in electron-beam tube. In a broadcasting station there is a master, a high-stable calibration oscillator, which transmits the information about its own frequency by using a television signal. Your television receives this information and a special device, a clock unit, tunes a generator- slave to the frequency of generator-master. The slave is not such high-stable as the master and all the time it is necessary to observe that it does not "sidetrack"its own frequency.

Such observation has need for feedback: as soon the slave begins to "side-track"its frequency as the clock unit compares master's and slave's frequencies and "enforces"the slave to operate again on master's frequency.

Thus, the "master-slave"principle involves a monitoring and feedback for each slave to be forced to execute its operation function.

Another example of dynamic planning and control by the master-slave principle is a conveyor. The master is a high-stable conveyor velocity, the slaves are workers, carrying out independently of each other the similar operations with the velocity that is "enforced"by the master. The nonperformance of necessary operations with a right velocity is detected at once (monitoring) and then is corrected (feedback).

A shining example of the "master-slave"principle is a pirate team. Without the hardly synchronized fulfillment of captain commands the sailing ship control is impossible. Therefore the absolutely free men, rovers, organize a "master-slaves"system, in which they discard for a while their freedom.

This hard principle of control in the people collective one tries usually to soften by some illusions of "social copartnership"and "corporate responsibility". However, in fact, a master remains a master and a slave is a slave.

6. The principle of continuous successful process

The principle of continuous successful process is also a dynamical control principle for stabilizing a system and preventing the occurrence and development of instability, which can lead to the chaos and collapse of system (from within).

In the mathematical theory of dynamical systems it is well known that the fore-runners of unstable processes are the oscillatory or "differently directed"motions. Therefore if in the dynamical system each of its subsystems evolves with positive derivative, then the instabilities, as a rule, are suppressed.

The above-mentioned principle is used, long ago, in a personnel policy of large west companies. In the case under consideration it is transformed in the following rule.

It is not recommended to demote a worker and to cut its salary. Only onward movement! To the point of the last day of the work in the company until he takes a letter of firing.

In other words: it is better to fire than to demote.

In this case the elementary subsystem, a concrete worker, moves upward continuously.

It may be remarked that the nosedive of any activity ratio can lead to a partial or complete destruction of all system much like the Great American depression in 30s of the last century: at the time the stock prices in Wall Street have fallen down suddenly and roughly.

7. The principle of wide latitude

The principle of wide latitude (the principle of federalism) is, in some measure, opposite to the "master-slave"one.

Obviously, such a hard and single-channel control as the master- slave principle does not accept for all cases. A hierarchical federative control system is often more adaptive for performance of preassigned requirements.

Here we also have a master (for example, the president) but together with the master there are apprentices (for example, the governors) to whose the master devolves some his credentials. These credentials must be rather wide since otherwise the system will be close to that, operated by the master-slave principle. In this case the work is carried out on the boundaries of stability ranges. Therefore it is necessary to provide for a series of special administrative untidisaster measures in the structure of system.

Firstly, in the considered system it must be covered the mechanism of a quick and resolute change of apprentice (and in practice it must be applied automatically) in the cases when the apprentice does not deal with incumbent incident and requirements.

Secondly, it is necessary to provide for a system of matching, training, and education of apprentices, what is also a stabilizing factor.

Obviously, in the case of the well-developed and correctly operational federative system, all the players obey a certain system of the established and conventional rules. In this case there occur some traditions and a system of values, which all participants of process are oriented on.

The above implies that the frequent and regular removability of master and apprentices is by no means a stabilizing factor, rather quite the reverse. In each of these alterations there are elements of destabilization. The more is a complicated controllable dynamical system, the greater is the probability that they can reveal themselves.

References

1. Leonov G.A. chaotic dynamic and classical stability theory. Regular and Chaotic Dynamic. Moscow. 2006 [in Russian]