Abstract
The study focuses on revealing the typical patterns of COVID-19 spread. The identification of regularities is based on conceptual processing data, in which a sample is approximated by model functions expressing a specific phenomenological concept of the process in question. The reconstruction of desired parameters is hindered by insufficient information on the growth features, the limited volume of the required experimental data, and incomplete elucidation of the phenomenological properties of the objects required to formulate complex models. The investigation addresses these limitations using the Verhulst equation, which expresses the general factors of object growth and adequately describes the growth dynamics in the context of significant theoretical uncertainties regarding an object under study. The need for identifiability and numerical stability of the model parameter estimates is also addressed, and a special regularization approach is applied to obtain an adequate and stable picture of the virus’s evolutionary dynamics in the context of significant uncertainties. The results demonstrate that the growth of the viral population can be integrally described by two variable parameters of the Verhulst equation. Interpreting the reconstructed functions indicates the existence of the typical pattern of seven stages of COVID-19 spread. The possibility of radically managing the number of infected persons is established. The proposed processing of coronavirus cases is of practical value for the preclinical diagnosis and control of virus spread in any geographic area, such as a district, city, local region, or country.

Key Words: COVID-19, Data Processing, Evolutionary Patterns, Mathematical Simulation, Population Growth

Introduction
Identifying the evolutionary patterns of COVID-19 is an urgent challenge for mathematical techniques and computational mathematics. Investigations of COVID-19 spread can be considered at different description levels, ranging from microbiological mechanisms of infection to the social consequences of the disease’s spread. All of these types of studies are now becoming widespread. In particular, a comparative molecular dynamics simulation was conducted to understand the possibility of inhibition at the atomic level; the viral genomic RNA extraction from culture lysate and clinical specimen was described in protease inhibitors under bioactive molecules were studied in a real-time PCR detection assay on the basis of the genome sequences was developed in the incubation period of COVID-19 and its public health implications were described in a detailed virological analysis of virus transmission and its replication in tissues of the upper respiratory tract were reported in the immune responses induced by SARS-CoV-2 infection were considered in a virus spread and transmissibility investigation using epidemiologic data on incubation periods and serial intervals between cases in transmission chains were studied in the pathogen, clinical features, diagnosis, and treatment of COVID-19, including epidemiology and pathology aspects based on current evidence, were reviewed in the conditions of COVID-19 transmission were described in organizational resilience during the COVID-19 pandemic was discussed in and a policies package for achieving measurable health outcomes was considered in [1-12]. A detailed review of publications on this topic is given in [13].

Recognizing the undoubted importance of developing a theory of COVID-19 propagation mechanisms, we also point out the need to identify the phenomenological features of the process in question through the direct processing of observations rather than theoretical simulation. Therefore, a crucial question is “What is the life cycle of the COVID-19 progression in a given geographical region?”. The COVID-19 life cycle, which reflects all scenarios of the virus activity in a given geographic region, has not yet received attention.

The answer can be obtained by the model-based (not statistical) processing of real observational data on COVID-19 cases. Note
that the theoretical simulation of virus spread through mathematical modeling with predetermined spread scenarios provides important information. However, the theoretical simulation cannot capture the full diversity of COVID-19 adaptation in the real environment. Processing real data is required to exclude the postulation of the principal properties (for example, constant or variable properties) of the object in question. The results of an approximation of the observations are a critical source of new information regarding the characteristics of COVID-19 spread. Processing becomes more powerful if a certain phenomenological concept and its mathematical model are involved in approximating a given sample [14]. Special regularization methods should be applied to take into account the peculiarities of processing discrete samples with a very high level of uncertainty about the original information [15]. All the necessary theoretical background for numerical regularization is described in [16].

Materials and an Object Under Study
The characteristics of COVID-19 progression are revealed based on coronavirus cases that have been presented as public data [17]. Four countries – Germany, Sweden, Switzerland and Brazil – are considered and their COVID-19 cases are processed. These countries are selected because they reflect the most characteristic and, at the same time, different strategies for protecting the public from COVID-19 and different methods to obtain, accumulate, and publish information. Of course, the number of existing combinations far exceeds this choice, and these will receive attention in the future. The first step in this direction is to evaluate the common and private sides of the spread of COVID-19 in substantially different countries and whether the spread of COVID-19 can be managed at all if the goal is to limit the number of infected persons.

The different dynamics of the disease in regions reflect the various conditions that are known to exist for the virus spread. Nevertheless, this does not preclude the existence of general patterns of infection evolution, despite different conditions. The confirmed number of infected persons and the rate of the disease spread reflect the evolutionary process that can be used to estimate the virus activity. This can be done by treating the observations with functions that express the law of population change over time.

Processing should take into account that both understanding the nature of the infestation and obtaining the necessary quantitative information on the spread of the infection are subject to considerable uncertainty. Therefore, the above-mentioned observations should be processed by means of a mathematical model that reflects the most significant properties of the object. The energy redistribution equation of the object can serve as such a fundamental model. From this point of view, many physics objects are very revealing. For example, consider an airplane that accelerates on the runway and gains altitude. An airplane is a very complex system that, during acceleration, must provide the kinetic energy needed to perform the lifting work. In other words, it must increase its potential energy. In this case, the system goal (airplane control) as a whole is to achieve the required takeoff speed and to manage its technical components (the thrust level, flap positions, etc.) during the lift when external conditions (the flight aerodynamics) change. Concurrently, a special part of the system (the pilot) decides how the system components should respond to ensure stable altitude gain under any changes in pressure, temperature, density, and other similar atmospheric properties. The trajectory of an airplane is a typical S-shaped curve, and the logistic curve is the optimal trajectory that satisfies the minimum principle of the system’s energy expenditure [18]. A deviation from the optimal trajectory is evidence of a change in external conditions, and a corresponding reaction of the system (by its relevant technical components) to these changes. To calculate the trajectory, of course, the speed and driving directions are needed. However, if we consider the inverse problem, setting a trajectory allows us to determine many important factors, for instance, the initial conditions of the mission.

As is known, the S-shaped curve is a solution to the Verhulst equation [19]. Accordingly, the logistic model essentially expresses the balance between the energy received by an object at a lower energy level and its state, after moving to a higher energy level [18]. From a biological point of view, it is a balance between individuals that appeared and disappeared [19]. In this case, population changes contain important information about the dynamics of the evolution in the form of an “environmental impact-object reaction” protocol.

Viewing the virus spread as a transition of a population from one state to another, changes in the S-shaped curve can be interpreted as the population response and adaptation to current impacts of the external environment. As a result, we have a source of information on whether the population has gained/lost enough energy to develop/decay. Consequently, the S-shaped curve holds information on all stages of activity/decay of the current virus modification. This information is presented as a finite sequence of strictly defined forms of increasing or decreasing parameters of the Verhulst equation, which determine the growth potential. The number of stages, of course, cannot be singular, but the processing should show that the population evolution under different conditions is either essentially individual, or that there is a finite set of typical evolutionary stages.

Thus, the objective of the study in this paper is to identify patterns in the dynamics of deviation from the classical S-shaped curve due to environmental exposure and the adaptation of the virus to it. Determining all the virus stages of activity/decline allows for further meaningful detailing of the mathematical model of viral progression, since the processing reveals factors affecting the virus evolution potential. Applying the model functions that reflect a certain conceptual framework in data processing provides an answer to the question “What should be done because?” (for example, sharp changes in the values of the factor in question mean that the processes occurring in a certain vicinity require further model detailing) as opposed to theoretical simulation that answers the traditional question “What happens if?” (that is, the parameters, input
Model and Methods

Let us process the public data by approximating the discrete sample with solutions to the Verhulst equation with variable parameters:

\[
\frac{dy}{dt} = a_1(t)y - a_2(t)y^2, \quad y|_{t=0} = y_0, \quad t > 0,
\]

where \( y_0 \) is an initial state, the function \( a_1(t) \) denotes the renewal factor of growth that in the general case should be time-dependent, and \( a_2(t) \) expresses the decline factor that is also time-dependent.

Model (1) reflects the main regularities of object adaptation in a given environment without introducing any environmental model based on a general balance within the framework of the "expended-reached" formulation. It should be noted that there are numerous phenomenological growth models reflecting the variety of factors that can act on the virus spread. More complex simulation of viral evolution is based on different kinetic schemes [19]. The introduction of additional factors into the mathematical model, which express the so-called susceptible-infected-recovered-dead schemes, will undoubtedly improve the simulation. For example, a model with thirteen factors was proposed to estimate the effectiveness of mask introduction [20]. However, their application requires postulating a number of properties and mechanisms of the object in question, which in the case of COVID-19 are still uncertain. From this viewpoint, the approximation of observed data by the model functions satisfied (1) adequately reflects changes in the number of infections under conditions of weak formalization of both the infection source and the viral habitat.

Despite the simplicity of the model, the Verhulst equation adequately reflects the behavior of the object as a whole because of its generalized character. Importantly, it takes into account two main factors, birth and death. Owing to them, the viral state – growth or degradation – is defined as the difference between the number of disappearance and appearance individuals. Accordingly, it is possible to introduce a characteristic such as a growth potential, which would reflect the current capacity of the viral population to change. The dynamics of change in potential show the directions of the object evolution toward growth or degradation.

It is important to note that the SIR-model transformations after linearizing the resulting nonlinear equation, reduce the SIR-model to the Verhulst equation [21, 22]. This emphasizes that the Verhulst equation provides an adequate and accurate simulation of the virus epidemic. Furthermore, considering the difficulties of simulation of habitats with social factors [23, 24], it can be claimed that the Verhulst equation, which integrally expresses the growth of any object in an external environment, can provide a productive interpretation of the action of any external factors. As a result, we can obtain a general protocol “impact – reaction” where basic regularities are revealed under significant modeling uncertainties.

Since the proposed approximation belongs to the class of ill-posed problems, its numerical implementation requires special measures [15]. For this purpose it is applied The approach [16]. It proposes identifying various representations of unknown parameters by regularizing the global and local numerical instabilities and does so in a series to ensure the stable reconstruction of a large approximating basis. Sequential solutions to inverse problems ensure the identifiability of desired parameters that belong to an invariant family. The locally sequential refinement restricts local spikes in addition to the general regularization under a scheme of separate matching with observations. In total, relaxation of a solution eliminates disturbances at individual points during a complicated function approximation. The desired parameters are reconstructed in the absence of information about their functional type and the presence of a sample whose size is smaller than the number of unknowns.

The desired parameters \( a_{1,2} \) of model (1) are approximated by the piecewise linear functions

\[
a_{1,2} = \frac{X_{i+1}^{(1,2)} - X_i^{(1,2)}}{\tau_{i+1} - \tau_i} t + \frac{\tau_{i+1}X_i^{(1,2)} - \tau_{i}X_{i+1}^{(1,2)}}{\tau_{i+1} - \tau_i}, \quad \tau_i \leq t \leq \tau_{i+1}, i = 1, N_{1,2} - 1,
\]

where \( \{\tau_i\}_{i=1}^{N_{1,2}} \) is the approximation grid, \( X = \{X_{i}^{(1,2)}\}_{i=1}^{N_{1,2}} \) denotes the nodes of the sought functions \( a_{1,2}(t) \) and \( N_{1,2} \) is the number of nodes. Consequently, the approximation basis is defined as \( \{\tau_i X_i^{(1,2)}\}_{i=1}^{N_{1,2}} \). The use of piecewise linear functions facilitates reconstructing very complex and beforehand unknown dependencies due to their high parametrization in terms of the number of nodes to be identified. The required regularization of the functions \( a_{1,2}(t) \) reconstruction is conducted in two separate schemes. The first is as follows:
\[ a_2^{(\delta)} = \text{Arg inf} \left\{ W_1 \int_0^T a_2^2 dt + W_2 \int_0^T a_2^2 dt + W_3 \int_0^T \left( \frac{d^2 a_2}{dt^2} \right)^2 dt \right\}, \]
\[
\sum_{i=1}^n \left[ y(t_i) |_{a_{1,2}} - y_i^{(\delta)} \right]^2 \leq \delta_1,
\]
\[
\int_0^T \left| \frac{dy}{dt} \right|_{a_{1,2}} - \frac{d\varphi}{dt} \left| dt \leq \delta_2, \right. \]

and the second one is as follows:

\[ a_1^{(\delta)} = \text{Arg inf} \left\{ W_4 \int_0^T a_1^2 dt + W_5 \int_0^T \left( \frac{da_1}{dt} \right)^2 dt + W_6 \int_0^T \left( \frac{d^2 a_1}{dt^2} \right)^2 dt \right\}, \]
\[
\sum_{i=1}^n \left[ y(t_i) |_{a_{1,2}} - y_i^{(\delta)} \right]^2 \leq \delta_1,
\]
\[
\int_0^T \left| \frac{dy}{dt} \right|_{a_{1,2}} - \frac{d\varphi}{dt} \left| dt \leq \delta_2, \right. \]

where \( y^{(\delta)} \) is a sample, \( y \) denotes the model function with the current values of the desired parameters \( a_{1,2} \), \( y \) is a spline approximation of the sample, \( T \) is the final time of the observations, \( \delta_{1,2} \) denotes a consistency corridor that is determined by regularization of a given sample, \( W_{1-6} \) denotes the weight coefficients.

The stabilizers in formulas (3) and (4) are presented for cases with a sufficient number of differentiable sought functions. In these, the first derivatives characterize the rate of change and the second derivatives reflect the smoothing degree of the functions being sought. The stabilizer minimum is a condition to exclude numerical oscillations [15]. Turning to discrete forms in problems (3) and (4), the sought function and its first derivative are approximated by the linear-polygonal function (2), and the second derivative is estimated by a finite-difference over three neighboring points. Due to this, the necessary account of the reconstructed function complexity in terms of its curvature, which should tend to a linear form, is accomplished. The minimization of such a discrete analogue in (3) and (4) reduces the accuracy requirements for the approximation of the second derivative. To restore essentially nonlinear dependences, the most critical requirement is to extinguish both global and local numerical oscillations. From this point of view, the use of a three-point finite-difference approximation is acceptable. The selection of the best stabilizers was considered in [25, 26].

Note that the matching with observations is performed on both the sample and its derivative. The latter is obtained by a smoothing spline [27]. The sole parameter, \( a_2 \) or \( a_1 \), is reconstructed with another fixed parameter, \( a_i \) or \( a_s \). The model function \( y \) is determined from Equation (1) numerically by the Kutta-Merson method for the current parameters \( a_{1,2} \). The solutions to the variational problems (3) and (4) are sought by the penalty function method and the coordinate descent algorithm.

**Results**

The processing on coronavirus cases that is based on model functions satisfied the Verhulst equation and the reconstruction of its two variable parameters demonstrate the high informativeness of the results obtained. The nature of the model parameters integrally reflects the processes involved in the spread of COVID-19 and characterizes the viral dynamics within the framework of the “impact – reaction” protocol. As generalized functions, these parameters encompass all the processes that occur during the adaptation of the virus to the environment and summarize the interaction and transport mechanisms.

The first parameter \( a_i \) can be defined as a renewal factor of viral population growth. This parameter determines the viral properties that are necessary to increase its population. The second parameter \( a_s \) reflects the difference between the number of deaths and the number of births of the virus at any moment. This parameter can be defined as a decline parameter of viral population growth. It reflects the restrictions of growth throughout viral life and demonstrates how the factors of viral decline prevail and dominate over the factors of its appearance. Figures 1–5 show the typical patterns of the desired model parameters. Countries with different protection strategies against COVID-19 were considered. They reflect radically different dynamics of COVID-19 spread. However, general patterns can be revealed.
Generally, for COVID-19 spread, the character of its renewal factor \( a \) after the initial stage can be defined as a piecewise constant function. At the initial stage, the renewal parameter has a completely pulsatile character (Figure 1). The decline parameter \( a \) demonstrates its oscillatory nature in the general and local sense around a certain curve (Figures 2–5). This parameter takes both negative and positive values. Negative values mean that the number of newborns dominates the number of deceased individuals. The most active stage of growth occurs when the number of appearing individuals exceeds the number of disappearing individuals, that is, when the value of the decline parameter is negative. Significantly, these stages were revealed in geographical regions with widely varying spread conditions and fundamentally different tactics for protecting the population.

The processing of coronavirus cases based on the conceptual model of virus spread shows the existence of seven stages of COVID-19 spread (see Figures 2–5). They constitute the life cycle of the virus spread. Each stage has its own morphology, which expresses a difference in the tendency to grow. These stages are completely repeated in the following waves of virus spread.

The first stage expresses the generation of a viral population. A characteristic feature of the stage is the existence of the highest renewal potential (the difference between the nearest maximum and minimum values). A vast area with negative values shows the initial renewal potential of the emerging population. This period corresponds to the most significant changes in the parameters of viral dynamics (see Figure 2a). At the same time, the increment of infected persons remains at a low level. Protective mechanisms against infection spread serve as a barrier function because the potential for disease growth remains significant, which manifests itself in the following stages. The first stage can be determined as the generation of a viral population.

The second stage reflects the period in which the growth rate of infected persons begins to increase. Here, the amplitude of the decline parameter increases sharply, having mainly negative values (see Figure 2b). The changes in the domination between the appearance–disappearance factors they have considerable frequency. This means that all the mechanisms of population growth have already been set in operation but have not yet enough power. It could be said that this is the most sensitive stage of growth because it shows whether or not the population can overcome the protective barrier of the environment. Here, the distinctive feature is a sequential decrease in the amplitude of oscillations to a specific value that weakly changes during the next period. The second stage can be determined as a speed-up of growth.

The third stage indicates the beginning of increased growth after the previous weak increment in the number of infected persons. A distinctive feature is the stable but slowly changing amplitude of the oscillations around a certain value. This means that the potential for the growth of the viral population demonstrates its continued presence for renewed infection (see Figure 2c). The numerical simulation of the second and third stages demonstrates their most instability of the calculations. This fact reflects a very active stage of virus adaptation to the environment. From the viewpoint of the rate growth of the coronavirus cases, this stage can be determined as a moderate increment of growth.

The fourth stage reflects the complete transition to exponential growth and defines the achievement of a local peak for the increase in coronavirus cases (see Figure 2c). A characteristic feature of this phase is the contraction of the oscillations of the decline parameter to the vicinity of the small value (see Figure 2d). This is an indication of how the decline parameter diminishes during the transition from negative to positive values. The shift to the positive region means that protective mechanisms against infection spread begin to manifest themselves fully and are transformed from a barrier function to a suppression one. The stage can be determined as an abrupt growth.

The fifth stage denotes the period of COVID-19 spread, in which the growth and decline factors begin to decrease and tend to a specific small value (see Figure 2d). The amplitude of the changes in the decline parameter is large locally. This means that the potential for viral population growth remains significant despite its decrease. The stage can be defined as a resistant renewal.

The sixth stage means the period when the variability of the decline factor is systematically reduced (see Figure 2d). The recurrent cases of coronavirus have steadily dropped, and the oscillations of the decline parameter tend to have a smaller amplitude. This fact suggests that factors of disappearance of individuals are beginning to play a major role. The stage can be determined as a regression of the renewal potential.

The seventh stage expresses the period of COVID-19 spread, in which the variability of the decline parameter is tiny, and its value tends to be constant. Zeroing the renewal potential means complete termination of the renewal and transition to a stable equilibrium "virus – environment". Latter is associated with a plateau on the S-shaped curve. Data processing shows the existence of two variants of growth potential reduction: from the area of positive values (Figures 2d and 4d) and from the area of negative values of the decline parameter (Figures 3c and 5d). This stage can be determined as attenuation of the renewal potential. The absence of a trend to zero in the behavior of the decline parameter means that there is a potential for the formation of a new wave of infection (Figure 2d).

The stages described above are repeated in four countries with different dynamics of COVID-19 prevalence and are presented in Figures 3–5. The results reflect that the infection after its outbreak evolves in a strictly defined sequence. Their duration and forms of manifestation appear to be individual, depending on the protective measures taken, but in all the cases considered, the morphology...
of viral evolution, that is, the periods of growth and decline of the spread, includes seven stages. Data processing shows that it is impossible to completely exclude the occurrence of any of the stages defined above from the full life cycle of COVID-19.

The results also indicate that all technologies that limit the virus spread can be divided into two general groups. The first group, for which the parameter \( a_1 \) is responsible, is the blockage of virus renewal conditions, namely, everything that involves the evolution of the virus infection at its microscopic and mesoscopic levels. These levels are not responsible for the interaction with the environment and form the potential for renewal. The second group, for which the parameter \( a_2 \) is responsible, is the generalized representation of COVID-19 adaptation under medicinal, physiotherapeutic, organizational, and other similar means that lead to the suppression of the viral spread. This group takes into account actions at the level of the external protection contour of the virus. Its overcoming is reflected in an increase of the parameter \( a_2 \).

From this point of view, mathematical modeling of the influence of these groups on the ability or inability to restrict the virus spread is of undoubted interest. An analysis of the properties of the solution to Equation (1) shows that at any point in time there is a possibility to decline the growth of the S-shaped curve down to a zero-infected person. Managing the number of infected persons is achieved through protective measures, as described above, fall into two groups. However, the implementation of such measures requires the fulfillment of radical conditions. First of all, Equation (1) shows that the condition \( a_1 \rightarrow 0 \) must be satisfied in order to meet the requirement of zero-infected persons. If \( a_1 \) remains at a level different from 0, then starting from some threshold value the total number of infected will increase again. The second condition to decrease the number of infected persons is to strengthen the suppression of virus development at all levels. To do this, two conditions must be met: (i) the decline parameter \( a_1 \) should have an abrupt change, and (ii) its subsequent tendency towards negative values must be majorized by the value of \( a_2 \). Data processing demonstrates (Figures 2d and 4c) that a similar managing by COVID-19 spread is acceptable. However, if the level reached by the decline parameter \( a_1 \) after its transformation (i) and recovery according to requirement (ii) becomes significantly lower than \( a_1 \), even in the positive values range, then the number of newly infected persons increases again. Figures 2d and 4c indicate just such cases. Note, however, that China’s protective measures against the spread of COVID-19 are implemented in such a way that the number of new infections stabilizes at a low level. This confirms that conditions (i) and (ii) can be met successfully in practice. The qualitative analysis presented above of the properties of the solutions to the Verhulst equation requires quantitative criteria for the effectiveness of limiting the spread of viruses. These criteria should be further investigated theoretically.

Since the present study is based on the processing of reported COVID-19 cases rather than on the analysis of theoretical evolutionary models, the degree of generality of the results obtained does not cover all possible variants. Nevertheless, the processing of observations provides key information for the development of theoretical models. And in this regard, the typical stages of COVID-19 spread and conditions of infection control found are essential modeling directions. The obtained results signify the importance of further research of the spread dynamics of COVID-19 to generalize outcomes.

The practical use of the results obtained can be aimed at assessing the effectiveness of the measures taken to protect against infection and to diagnose the current state of the virus spread according to its life cycle.

**Conclusions**

Mathematical simulation of COVID-19 spread by solutions to the Verhulst equation with variable parameters provides practical and significant information on the current state of the disease and its potential for further spread. The morphology of viral evolution, the potentials for COVID-19 spread, as well as the features of viral propagation are expressed by the reconstructed renewal and decline factors. The essence of these factors is established by the volumes of appearing and disappearing individuals.

Data processing and interpretation of the obtained results indicate the existence of the typical pattern of seven stages of COVID-19 spread. This number persists despite the significantly different conditions that exist and defense tactics that have been used against virus propagation in the four countries examined. The stages constitute the life cycle of COVID-19. Knowing this life cycle facilitates the recognition of the actual state of the pandemic and is valuable in diagnosing its stage of progression.

It was determined the possibility of managing the number of infected persons down to a zero-infected number through specific protective measures at each stage of the COVID-19 life cycle.

**Acknowledgements**

The author would like to thank the reviewers, whose questions and comments have improved the presentation of the results and have supported the publication.

**References**


Figure 1: Reconstruction of the renewal factor of COVID-19 spread: a) and b) – Germany, c) – Switzerland.

Figure 2: Reconstruction of the decline factor of COVID-19 spread in Germany
Figure 3: Reconstruction of the decline factor of COVID-19 spread in Switzerland

Figure 4: Reconstruction of the decline factor of COVID-19 spread in Sweden
Figure 5: Reconstruction of the decline factor of COVID-19 spread in Brazil