

А.И. Леднев

**ПРЕИМУЩЕСТВА И НЕДОСТАТКИ АНАЛИТИЧЕСКОГО
(ФОРМУЛЬНОГО) ПРЕДСТАВЛЕНИЯ
ГЕОПРОСТРАНСТВЕННЫХ ДАННЫХ.
ВЫВОДЫ И СЛЕДСТВИЯ**

A.I. Lednev

**ADVANTAGES AND DISADVANTAGES OF ANALYTICAL
(FORMULA) REPRESENTATION OF GEOSPATIAL DATA.
SUMMARY AND EFFECTS**

Предложено аналитическое (формульное) представление пространственных данных, которое может явиться тем «ключиком», открывающим дверь на качественно новый уровень моделирования природных объектов и процессов.

Ключевые слова: представление, геопропространственные данные, информатика.

Analytical (formula) representation of spatial data is proposed herein, such representation may be the very «key», which opens the door to a new quality level of modeling of natural objects and processes.

Key words: representation, geospatial data, information science.

Введение

Тема статьи связана с работами кафедры гидрогеологии и геодезии РГГМУ, проводившимися ею в 70-х годах ушедшего столетия в зал. Терпения (о. Сахалин). Одно из направлений моделирования сезонных изменений морфологии дна залива было определено разработками В.С. Файна [1975], успешно внедренными в криминалистике и дизайне. Однако в те годы для зал. Терпения такое моделирование реализовать не удалось. Мне было предложено проанализировать такой подход на основе современных достижений информатики.

Представление (или модели) геопропространственных данных – способ цифрового описания пространственных объектов Земли, тип структуры пространственных данных. "Геопропространственные данные" означают информацию, которая идентифицирует географическое местоположение и свойства естественных или искусственно созданных объектов, а также их границ на Земле.

«В науках о Земле, имеющих дело с пространственными данными, сложилась ситуация, когда информационный «взрыв» соседствует с информационным «голодом». Одни специалисты сетуют на ограниченность сведений, что ведет к упрощению описаний, другие не успевают «перелопатить» горы материала» («Основы геоинформатики» под ред. В.С. Тикунова, 2004).» Появление геоинформационных систем произвело переворот в инструментарии моделирования географического пространства, реализовав способ его описания и представления в форме цифровых моделей. Моделированию вначале приписывали неограниченные возможности, особенно в области имитации сложных явлений во

всем их многообразии. Но именно их сложность и многообразие явились непреодолимым препятствием для разработки стандартных вычислительных программ и типовых схем, способных адекватно описать природное явление без искажения при приближении под созданные модели анализа.

В такой ситуации самое разумное не пытаться заходить глубже в тупик, решая данную ситуацию путем подгонки (расширения и упрощения) традиционных представлений геоданных под модели процессов (явлений), а попытаться найти принципиально новые способы и структуры описания, способные обеспечить прорыв в устранении имеющихся противоречий.

Существующие модели (представления) геопространственных данных

Цифровые пространственные данные не могут восприниматься человеком визуально или с помощью иных чувств, что создает необходимость их визуализации или иного чувственного восприятия. Поэтому и существующие в настоящее время способы цифрового представления (моделей) геопространственных данных исходят из возможности их последующей возможности чувственного восприятия (традиционно это карты).

При выполнении рисунков на компьютере или использовании уже готовых изображений обычно применяется два вида компьютерной графики: растровая и векторная. При этом обычно утверждается, что именно при помощи аппарата векторной и матричной алгебры вполне удовлетворительно описывается механизм работы зрения, особенно фасеточного глаза, например такого, как у стрекозы. Все основные вопросы с математическим аппаратом для описания произвольного образа оказывается, уже решены. Осталось немного улучшить программное обеспечение и использовать компьютер помощнее. Причем как при использовании векторной, так и при использовании растровой графики изображение на экране монитора формируется точками (пикселями). Различия между этими двумя видами графики главным образом состоит в способе записи информации об этих точках.

Векторная графика для представления графической информации использует совокупность линий или простых (подчеркнем: простых) фигур, описываемых математическими формулами. Все точки изображения рассчитываются непосредственно перед выводом по этим математическим формулам. Общее количество точек изображения также изменяется в зависимости от размера изображения и возможностей устройств вывода.

В отличие от векторных изображений при создании растровых изображений математические формулы не используются. Компьютерное растровое изображение представляется в виде прямоугольной матрицы, каждая ячейка которой представлена цветной точкой. Таким образом, растровое изображение чем-то напоминает мозаику. При оцифровке растрового изображения вся информация о каждой точке изображения записывается в виде набора чисел, помещен-

ных в матрицу (таблицу). Из-за того что информацию при использовании растровой графики приходится хранить в виде строго определенной матрицы (таблицы), файлы этой графики имеют большой объем. В матрице приходится хранить даже информацию о каждой точке фона. Затруднена работа с отдельными объектами изображения. Трансформация (перемещения, особенно с поворотами, масштабирование, деформации) обычно связана с искажением изображения. Несмотря на эти недостатки, растровая графика все же имеет наибольшее распространение благодаря своим главным преимуществам, заключающимся в технической реализуемости получения фотореалистичных изображений и автоматического ввода (оцифровки) графической информации при помощи сканеров, цифровых видео- и фотокамер, графических планшетов.

Различные способы записи информации о точках компьютерной графики и определили наиболее употребительные представления геопространственных данных: векторное, растровое, регулярно-ячеистое и квадратомирическое. Необходимо отметить, что они работают для объектов, размерности не более двух (планиметрических). Для представления поверхностей (рельефов) и других трехмерных объектов необходимо расширение представленных моделей в виде объемных пикселей – «вокселей» (трехмерное расширение растрового представления) и тетраэдров или октаэдров (трехмерное расширение полигональной векторной модели).

Таким образом, математической основой моделирования геопространства является аппарат векторной и матричной алгебры, который накладывает определенные ограничения в виде необходимости задания узлов не обязательно регулярной сети или системы «плоских (двумерных)» кривых с присвоением этому узлу каких-либо качественных параметров.

Возможности использования существующих математических моделей для представления (моделирования) природных процессов

При исследовании любого процесса любой природы всегда в первую очередь стремятся установить его описывающие функциональные зависимости, т.е. выявить и описать в формальном виде связи между величинами, характеризующими развитие процесса. При описании изучаемых процессов для обеспечения использования в дальнейшем результатов исследований необходимо задать функциональные зависимости (функции), проявляющиеся в этих процессах и характеризующие их. Формальное, прежде всего аналитическое, описание этих связей предполагает обязательное использование математических методов. При этом известно, что никакая математическая схема не в состоянии полностью во всех подробностях адекватно описать действительные явления, происходящие в ходе развития процесса. Поэтому традиционно в ходе исследования различных процессов осуществляется выявление основных или существенных особенностей наблюдаемых явлений, их схематизация или идеализация и последующее описание при помощи математических методов, чаще всего предпо-

лагающих использование относительно простых математических моделей. Однако современные требования, предъявляемые к полноте, достоверности и точности исследования процессов практически во всех областях знаний, вызывают стремление к более детальному изучению и соответственно описанию явлений, что приводит к быстрому усложнению, используемых для этого математических моделей. В качестве процесса рассматривается последовательная смена состояний в развитии наблюдаемых в природе явлений. При этом состояния характеризуются переменными величинами, а сами эти величины изменяются с течением времени или в зависимости от изменения других величин в соответствии с зависимостями, которые в ходе развития процесса сами претерпевают изменения, включая и скачкообразные. При этом одной из основных проблем математического моделирования этих процессов является непропорционально быстрый рост сложности и объемов, используемых для этих целей математических моделей, что все чаще приводит к получению математических моделей малопригодных для практического использования. Причем не улучшает ситуацию даже значительный прогресс вычислительной техники.

Вероятно, следует критически взглянуть на основы математического аппарата представления «геопространственных» данных. Аппарат векторной и матричной алгебры рассматривался в качестве основного математического аппарата представления потому, что он способен описывать произвольную точку и соответственно любую сумму точек в пределах выбранной матрицы. Однако необходим математический аппарат, способный описывать произвольную точку и произвольную сумму таких точек не в матрице, а в произвольном пространстве. Свойства выбранной матрицы оказывают решающее влияние на вид представления данных. Необходимо также учитывать, что для совершения последующих операций с полученными матрицами они должны иметь сопоставимыми свои соответствующие параметры. Используя аппарат векторной и матричной алгебры, мы описали не только образ точки, но еще и "фон" (всю площадь матрицы). Причем если мы хотим получить описания, сопоставимые друг с другом, как бы мы не старались, оторвать образ точки от "фона" не удастся, так как это связано с коренными свойствами аппарата векторной и матричной алгебры.

Состояние аналитического (формульного) способа представления пространственных данных

Основной задачей представления (моделей) геопространственных данных является изучение и описание природных объектов и связанных с ними процессов и явлений с помощью пространственных соотношений между ними. Описания производятся с помощью математических моделей, а процессы устанавливаются от известных физических закономерностей. Чем точнее «модель» соотносится с реальным объектом природы и вернее установлены функциональные зависимости, тем больше соответствует реальности данное представление. Чем

сложнее объект и больше связей с другими, тем дальше от реальности его «модель» или она настолько «громоздка», что не имеет однозначного решения, а значит и не имеет практического применения. Успехи геоанализа «технических» систем определяются соответствием их математических моделей описания и значительной степенью, зачастую искусственно созданной их изолированности. Природные объекты трудно поддаются формализации из-за размытости границ, фрагментарности положения и строения, хаотичности поведения, скачкообразности и прерывистости порождаемых ими или влияющих на них процессов. Если для математического описания технических систем еще достаточно аппарата векторной и матричной алгебры, так как они «вписываются» в пределы выбранной матрицы, то для природных систем необходимо делать описание произвольных «точек», состояния которых характеризуются переменными величинами в произвольном, зачастую многопараметрическом пространстве.

Если состояния «точек» описываются переменными величинами, то их значения можно связать функциональной зависимостью (каждому значению одной величины соответствует значение других величин). Из всех способов задания функций задание формулой наиболее предпочтительно, так как значение любой величины не зависит от способа и свойств задания (например, от размерности таблицы или масштаба карты). Поэтому понятно желание создать представление (модель) данных, величины которых «связаны» между собой при помощи формул.

Существующие математические модели реальных объектов идеализированы в разностной степени приближения. При более детальном описании всей модели происходит ее неприемлемое усложнение, при детализации ее части велика вероятность ситуации персонажей басни Крылова «Лебедь, рак и щука» или притчи о четырех слепых мудрецах. Ситуация осложняется тем, что процесс детализации любого реального объекта по существу является бесконечным. По мере детализации, приходится учитывать все более мелкие подробности и, в конце концов, любой объект перестает быть сплошным – рассыпается на отдельные составляющие его элементы, которые в свою очередь делятся на элементы, их составляющие, и т. д. Таким образом, реальные объекты являются дискретными, и в силу этого связанные с ними процессы, явления также дискретны. Подавляющее же количество существующих математических моделей представлено гладкими непрерывными функциями, а для описания дискретных объектов необходимы функции, допускающие скачкообразное изменение своих значений. Их приходится задавать, используя способ задания функций словесной формулировкой или системой уравнений, объединенных условиями вида "если ..., то ..." или "при ...", что создает определенные неудобства, обусловленные главным образом тем, что значения таких функций невозможно определить только при помощи вычислительных операций. Кроме того, «если мы считаем, что материя обладает бесконечно зернистой структурой, то возможность

применять к реальности строгое математическое понятие непрерывности сводится почти на нет» (Б. Мандельброт). Все же «при определенном масштабе и определенных условиях, можно использовать непрерывные функции» (там же). Только осталось определить масштаб и условия. Мы хорошо научились моделировать объекты в размерах Земли и на атомно-молекулярном уровне. А что же делать с пограничными «размытыми (нечеткими) множествами» и их стохастическими моделями?

Нужен такой математический аппарат, основания которого способны преодолеть пропасть между областью дискретного и областью непрерывного, между арифметикой и геометрией, «между дискретной, качественной, индивидуальной природой числа в «комбинаторном» мире счета (арифметика) и непрерывной, количественной, однородной природой пространства в «аналитическом» мире измерения (геометрия)» (А. Френкель и И. Бар-Хиллел «Основания теории множеств»). И хотя такой аппарат математиками еще не создан, ничто не мешает использовать его уже существующие основания для аналитического (формульного) представления геопространственных данных.

Идейная основа и инструмент реализации аналитического (формульного) представления данных

Базовым понятием в математике и физике является точка. Примем как данность, что «точка есть абсолютное пространство, имеющее положение, или точка есть размерное абсолютное пространство, то есть ограничено по размеру, но не по количеству» (Е.Б. Чижев, 2001), часть пространства, которое нельзя поделить, но можно ограничить числом. «Пространство точек есть логически мыслимая качественная субстанция, служащая средой, в которой осуществляются количественные протяжения и характеристики тех или иных относительных пространств». А «пространство чисел есть логически мыслимая количественная субстанция, ограниченная по количеству, служащая для измерения и пересчета качественного протяжения и характеристик тех или иных относительных пространств». «В пространстве чисел существуют только целые положительные и отрицательные числа», которые все же конечны, и их границей служит пространство точки. А дробные и трансцендентные числа «возникают только в результате взаимосвязи относительных пространств, которые имеют как количественные, так и качественные характеристики» (там же).

Формула (функция), с помощью которой можно задать точку, используя выше приведенные высказывания (идеи), уже с успехом применяются для аналитического описания скачкообразно и качественно меняющихся процессов, задания ступенчатых и импульсных функций, получения уравнений «чудовищных» фигур и поверхностей, а с помощью уравнений фракталов и тел (рис. 1 и 2).

Для демонстрации возможностей использования этого универсального уравнения лучше всего обратиться к его автору (А.М. Белов, 2000).

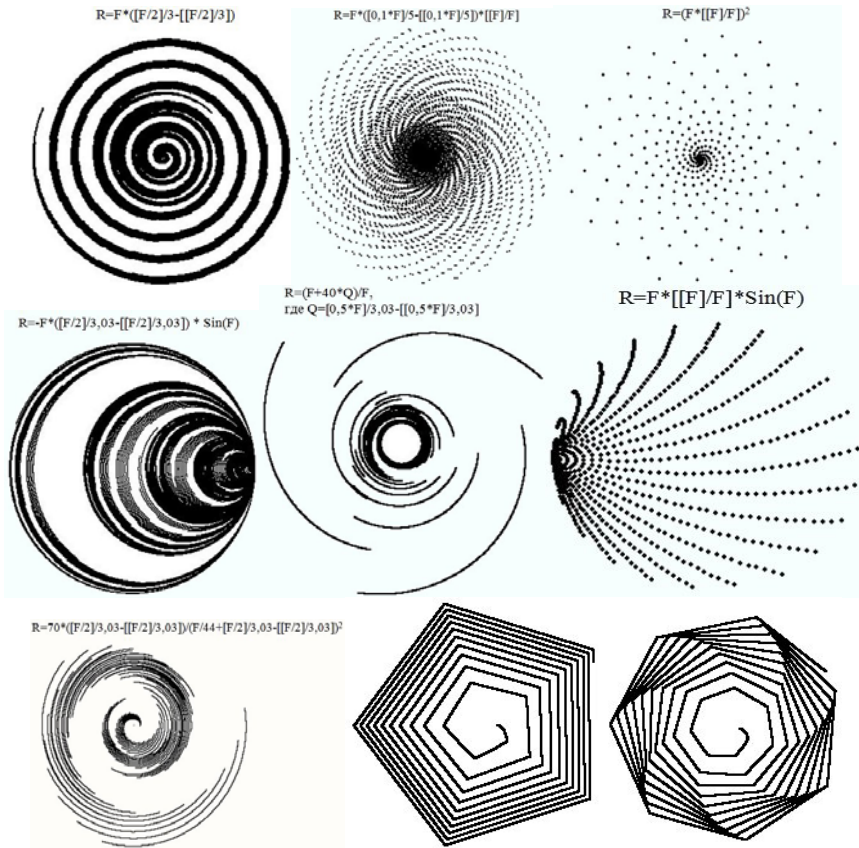


Рис. 1. Демонстрация возможности представления «немыслимых» фигур в полярных координатах с помощью «универсального уравнения» Белова А.М. (2000 г.)

Реализация аналитического способа задания функций этого типа была осуществлена за счет применения специального уравнения (1).

$$y = [x/x_i] \cdot [x_i/x], \tag{1}$$

где [] – знак, обозначающий целую часть числа; x_i – любое число.

Замечание: Знак [] в выражении (1) и далее по тексту статьи предполагает выполнение процедуры по отбрасыванию дробной части числа. Именно эту операцию обычно выполняют все компьютерные математические программы.

Функция, заданная уравнением (1), определена везде кроме точки $x = 0$ и $y = 0$ для всех значений x , кроме $x = x_i$ и $-x_i$, где $y = 1$.

Если уравнение (1) разделить само на себя, то будет получено уравнение (2):

$$y = ([x/x_i] [x_i/x]) / ([x/x_i] [x_i/x]). \tag{2}$$

График уравнения (2) состоит из двух точек $(x_i, 1)$ и $(-x_i, 1)$. Для всех иных значений x функция (2) значений не имеет из-за возникновения ситуаций деле-

ния на ноль. Проводить сокращения в уравнении (2) нельзя, так как тогда у всегда будет равно 1 и ситуаций деления на ноль возникать не будет.

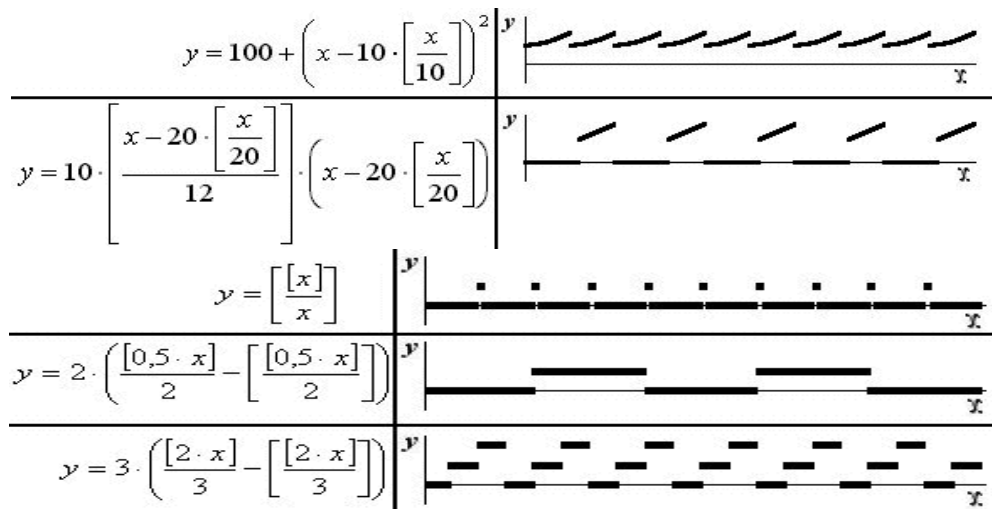


Рис. 2. Демонстрация возможности задания периодических, импульсных, прерывистых функций с помощью «универсального уравнения» А.М. Белова [1]

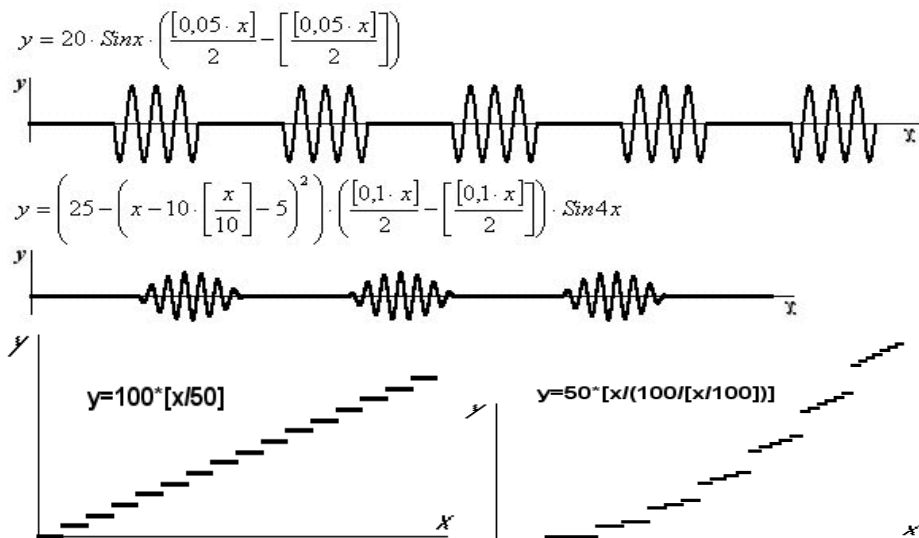


Рис. 2. Демонстрация возможности задания периодических, импульсных, прерывистых функций с помощью «универсального уравнения» А.М. Белова [1] (продолжение)

Исходя из свойств уравнений (1) и (2) можно сделать вывод, что путем изменения значений x_i или замены их в уравнениях различными функциями, а также комбинации уравнений (1) и (2) можно составить уравнения, задающие функции практически любой формы (см. рис. 1 и 2).

Универсальное уравнение в своем элементарном ($[x/x_i] \cdot [x_i/x]$) виде на графике задает точку (непрерывное изменение x ограничено числом x_i), а сумма этих уравнений – соответственно сумму точек.

Увеличив количество переменных, можно задать точку в трехмерном и более пространстве с заданными качественными характеристиками (применительно к графике задать толщину и цвет).

Если взять любую фигуру и начать бесконечно уменьшать однотипные части, её составляющие, то какими бы сложными эти части ни были, они постепенно все вырождаются в точки. Таким образом, можно говорить о том, что любая фигура должна графически представлять собой набор точек и может задаваться суммой универсальных уравнений. Причем суммы таких уравнений в общем случае будут бесконечными, так как при представлении какого-либо объекта точек теоретически к уже имеющимся точкам, составляющим объект, можно прибавлять бесконечное число новых точек с соответствующим увеличением густоты (плотности) размещения точек в пространстве. Трехмерные построения, выполняемые с помощью функций на основе универсальных уравнений, отличаются от обычной 3D графики. Обычная 3D графика обеспечивает лишь построение поверхностей объектов, а формулы на основе универсальных уравнений, весь объем, занимаемый объектом, включая его внутренности (рис. 3). Это обстоятельство очень важно, если их предполагается в дальнейшем использовать для выполнения с ними каких-либо операций, например расчетов.

Таким образом, формульное представление пространственных данных позволяет решить ряд проблем традиционного представления:

а) при работе с графическими источниками пространственных данных:

– поскольку в формуле осуществляется суммирование однотипных выражений, то порядковый номер точки изображения и место ее расположения в изображении (координаты) жестко друг с другом не связаны и могут соответствовать друг другу случайным образом, что удобно для автоматического ввода (оцифровки) и вывода графической информации при помощи технических средств;

– количество слагаемых принципиально ничем не ограничено, что позволяет увеличить количество описываемых параметров изображения и соответственно повышать качество изображений, добиваться более высокой их реалистичности;

– так как формула способна описать все без исключения графические объекты, как по отдельности, так и в сумме, то в программах в явном виде формулу использовать не обязательно, а в файлах достаточно будет сохранять лишь коэффициенты формулы;

– фон изображения может быть задан всего одной точкой, т.е. с использованием одного слагаемого формулы, что в значительной степени сокращает объем файлов, хранящих информацию об изображениях;

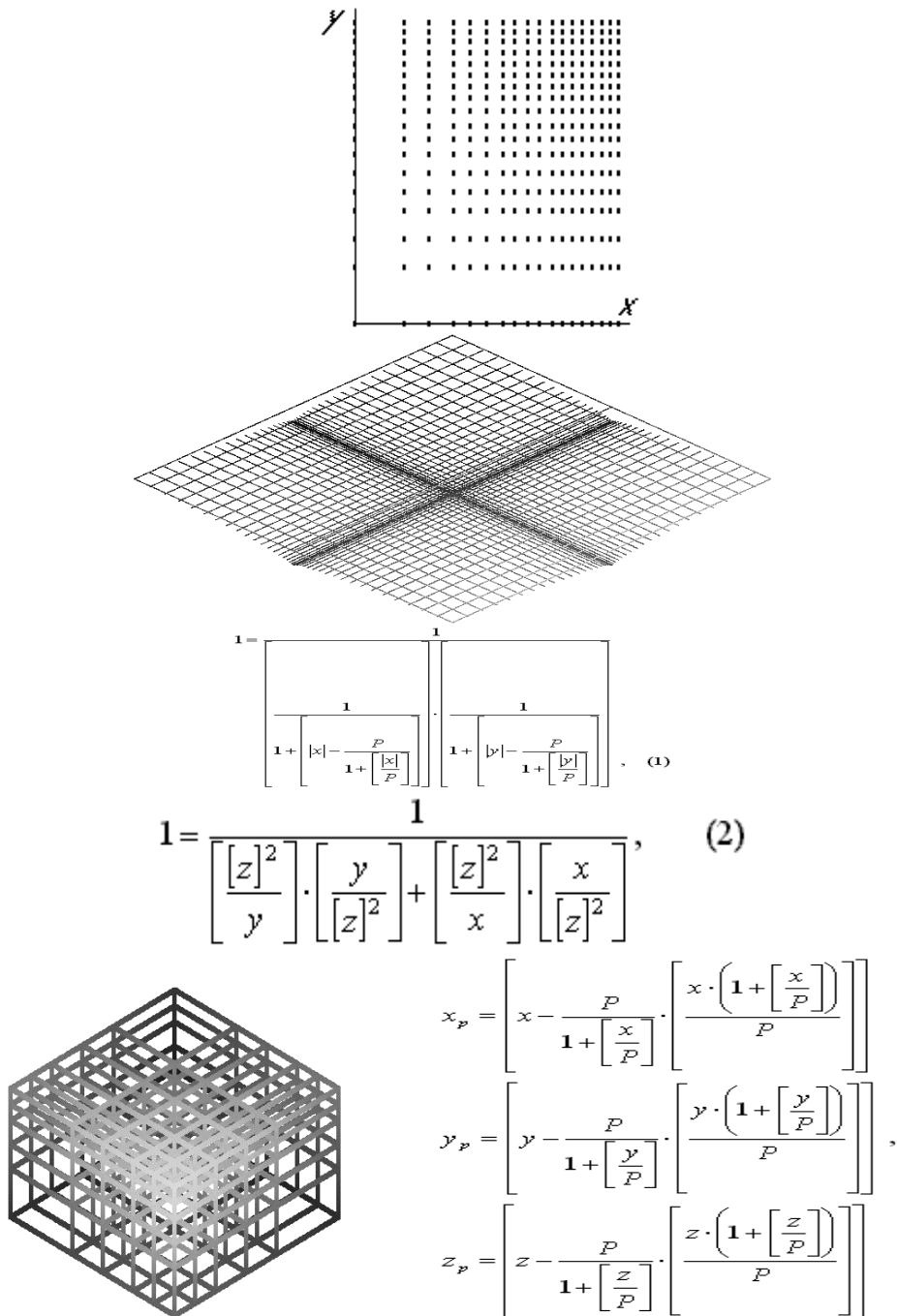


Рис. 3. Демонстрация возможности заполнения пространства точками плоскости(1), трехмерной плоскости(2) и объема(3) с помощью «универсального уравнения» А.М. Белова (2000 г.)

– формула обеспечивает описание изображения с возможностью полного или частичного перекрытия друг друга точек, формирующих изображение, что будет способствовать повышению устойчивости изображений к искажениям при их трансформации (рис. 4);

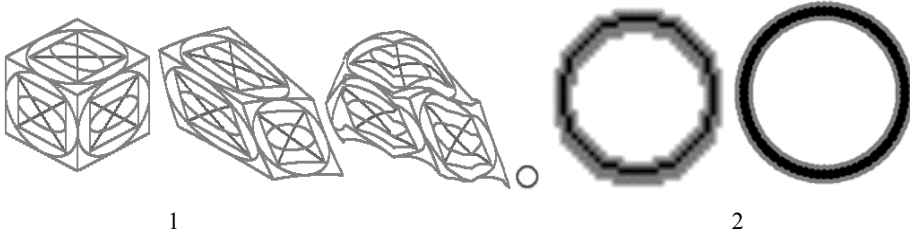


Рис. 4. Демонстрация устойчивости изображений аналитического (формульного) представления при деформациях и перемещении (1) или масштабировании (2)

– формула обеспечивает выполнение любых действий с любым коэффициентом формулы, что обеспечивает хорошие возможности по внесению изменений в изображения вплоть до изменения одной единственной выбранной точки изображения независимо от остальных точек изображения;

– так как суммируемые в формуле выражения не зависят друг от друга, то в состав формулы всегда можно ввести дополнительные слагаемые, что создает хорошие условия для работы различных алгоритмов по изменению изображений, например повышения четкости и контрастности изображений за счет введения дополнительных точек или увеличения (уменьшения) толщины (жирности) точек, формирующих изображение;

б) при работе с данными станций непрерывного наблюдения:

– с помощью универсального уравнения можно задать функцию целиком одним уравнением и выводить информацию в цифровом виде в любом интервале значений абсцисс или ординат;

– если информация о наблюдениях записывается в цифровом виде, с помощью программы можно оперативно выдавать ее в виде графика; и наоборот, непрерывную запись параметров можно оперативно переводить в запись в виде формулы (рис. 5);

– так как можно совершать операции над функциями и даже вычислять значения функции от функции переменной, то нет принципиальных проблем для одновременного измерения двух и более зависимых друг от друга параметров;

– формула (без искажения информации) может содержать интервалы значений параметра, для которых значения других параметров не определены;

в) при анализе и моделировании природных процессов и явлений:

– так как оцифровка изображения с помощью формул описания и преобразования при создании соответствующих программ (которые просто значительно ускорят процесс), вполне реально осуществима, можно соответственно описывать любой образ без «фона» и формирования матрицы с помощью суммы уравнений;

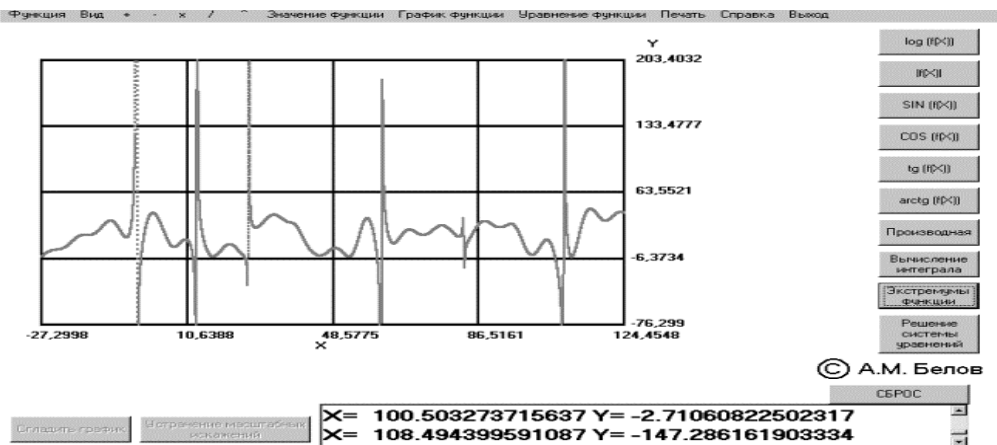


Рис. 5. Демонстрация возможности приближения функций, интерполяции, аппроксимации и составления графиков с помощью «универсального уравнения» А.М. Белова [1]

– так как уравнения зависят исключительно от свойств образа, причем даже можно использовать несколько свойств, процедура распознавания образов по заданным свойствам упрощается и на нее не влияет ни объем каждого из двух сравниваемых уравнений, ни размер, ни расположение образа, а сами уравнения легко преобразуются при изменении системы отсчета;

– преимущества формульного представления в виде устойчивости модели к масштабированию и трансформации (при перемещении как целого образа, так и его частей), возможности производства расчетов независимо от параметров «матрицы», позволяют моделировать природные явления в многопараметрическом пространстве с последующей его «визуализацией» в привычной нам плоскости (в виде сечения). Существующие в настоящее время представления (модели) геопространственных данных, основанные на векторной и матричной алгебре, вынуждены отталкиваться от обратных предпосылок – «из жестких плоских железобетонных плит собрать округлый изящный храм природы».

Выводы

Возможности современных представлений (моделей) пространственных данных, созданных с помощью аппарата матричной и векторной алгебры, вероятно, достигли своего предела. Увеличение мощности вычислительных средств и «расширения» программных продуктов только отодвигает по времени осознание этого факта. Нужен качественно новый подход.

Предложенное аналитическое (формульное) представление пространственных данных может явиться тем «ключиком», открывающим дверь на качественно новый уровень моделирования природных объектов. Залогом правильности выбранного направления являются: идейные основы геометризации основных понятий философии, физики и химии, заложенные в работах Е.Б Чижова, соответствующий им инструмент «универсального уравнения» А.М. Белова и

значительно расширяющий возможности фрактального анализа М. Мандельброта, а также «зернистость» строения и «гладкость» образованных природными объектами форм. Этим докладом я хотел показать только возможности принципиально нового подхода к анализу и моделированию природных объектов и процессов. И, может быть, уже настало время, когда, согласно «произвольному, но не противоречивому предположению» Бенуа Б. Мандельброта [2002], «наверняка мы столкнемся с такими случаями, для описания которых окажется проще использовать недифференцируемые функции, нежели те, что имеют производную. Когда такое произойдет, практическая ценность иррегулярных континуумов станет очевидной всем. Однако это – всего лишь мечтания. Пока».

Литература

1. *Белов А.М.* Функция для описания скачкообразно и качественно изменяющихся процессов // Тульские ученые накануне третьего тысячелетия: Сб. аналит. и информ. материалов. – Тула: Гриф и К^о, 2000.
2. *Бенуа Б.* Мандельброт Фрактальная геометрия природы / Пер. с англ. А.Р. Логунова. – М.: Институт компьютерных исследований, 2002.
3. *Чижов Е.Б.* Пространства. – М.: Новый Центр, 2001. или
Введение в философию математических пространств. – М.: Едиторал УРСС, 2004.
4. *Файн В.С.* Алгоритмическое моделирование формообразования. – М., 1975. – 142 с.