L-system modeling of neurons

B. H. McCormick

Scientific Visualization Laboratory
Department of Computer Science
Texas A&M University

K. Mulchandani

Scientific Visualization Laboratory
Department of Computer Science
Texas A&M University

June 14, 1994

ABSTRACT

A formal representation of neuron morphology, adequate for the geometric modeling of manually-traced neurons, is presented. The concept of a stochastic L-system is then introduced and the critical distribution functions governing the stochastic generation of dendritic and axonal trees are defined. Experiments with various stochastic L-system models for pyramidal, motoneuron, and Purkinje cells are reported which generate synthetic neurons with promising proximity to neurons in the neurobiology literature. Work is in progress to improve this degree of proximity, but more importantly to validate the derived stochastic models against available databases of manually-traced neurons. To this end a neuron morphology modeler is described which provides a methodology for iterative refinement of the stochastic L-system model.

1 RESEARCH SUMMARY

Two specific results of this research are methods to:

1. Generate a representative population of several neuron types (pyramidal, motoneuron, and Purkinje cells) from stochastic L-system models of these neuron morphologies; and
2. Validate the derived models against available databases of manually-traced cells.

Methods to parallelize these L-system models to grow graphical models of brain nuclei and other neural tissue are discussed.

2 MOTIVATION

L-systems were first proposed by Aristid Lindenmayer in 1968 as a mathematical theory to model the growth and morphology of simple plants.[2,6] Turtle geometric interpretation of the strings generated by the L-system grammar allows the geometric modeling of plants. Today L-systems are an accepted tool in plant modeling. Neurons, scaled by 10^5, show a strong resemblance to trees. This similarity led us to harness the power of L-systems to produce geometric models of neurons. Though neurons and glial cells, unlike plants, are composed of a single cell, analogous formal L-system methods are described below to generate neuron-like structures using separate subgrammars for the dendritic and axonal processes.

All models in this work are presently limited to modeling neuron morphologies at the limit of optical resolution. In particular spines are not presently included, except in a superficial way. Furthermore, fiber tracts and fiber terminations on cells are not presently modeled.

The applications of these L-system models are severalfold. First, the L-system morphological models bring

-8194-1698-3/94/$6.00

coherence to existing databases of neurons. They allow us to estimate the new information contributed by additional measurements of the neuron morphology.

Secondly, neurons are packed in a volume-filling manner into neural tissue, posing a barrier to visualization and interpretation of neural tissue scanned at the limit of optical resolution. New developments in confocal microscopy and multi-terabyte mass storage offer the potential for massively parallel tracing of fibers and other neuron processes in neural tissue, even at the scale of an entire rat brain. The interpretation of these massive volumetric data sets will rest on having creditable models of neuron morphology. Furthermore, because of the close juxtaposition of neurons, existing volume visualization methods, such as the marching cubes algorithm or the Levoy algorithm, would portray only a dense and confused volume. What is needed are object-oriented models for these volumetric databases of neural circuitry.

Finally, volume-filling tissue, because of interference between neighboring cell processes, can not be assembled from a collection of independently-generated neurons and glial cells. We believe the only creditable way to graphically generate morphological models of brain tissue is to mimic the growth processes of developmental neurobiology.

3 REPRESENTATION OF NEURON MORPHOLOGY

3.1 Terminology for neuron morphology

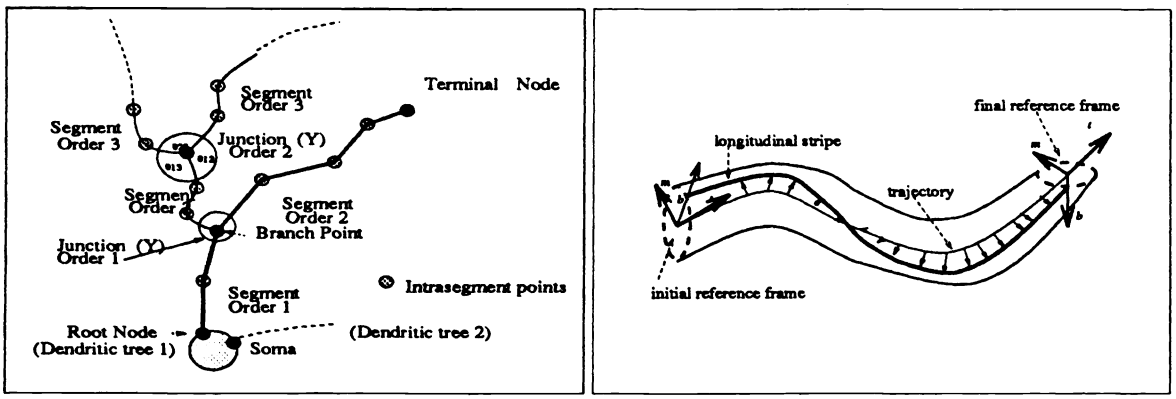
The constituents of a neuron are its soma, one or more dendritic trees (also called dendritic arbors), and its axon or axonal tree.(Figure 1 *left*)

Given a dendritic tree, three types of nodes are identified: (1) the *soma node* representing the point of attachment of the tree to the soma; (2) *branch points* at which the dendrite branches; and (3) *terminal nodes*, the endpoints of the tree. Starting at the soma node a unique path exists to any other node of the dendritic arbor.

A *dendritic segment* is the portion between two consecutive nodes in the dendritic tree. The *stem segment* has its origin at the soma node and is counted as the first order segment. Daughter segments arising from the first order branch point are called second order segments, and so on. A segment ending in a terminal node is called a *terminal segment*; all others are called *non-terminal segments.*

The *junction* at a branch point is called a *bifurcation* if the parent segment branches into two daughter segments, and a *multifurcation* if the junction gives rise to more than two daughter segments.

Axonal processes, including axonal arbors if present, are described by nodes, segments, and junctions in like manner, as discussed above for dendritic trees.



Left: Dendritic tree representation. *Right:* Segment representation

Figure 1: Representation of neuron morphology

3.2 Representation of segments

3.2.1 Segments are represented as generalized cylinders

Segments are represented by generalized cylinders (Figure 1 *right*) whose central trajectory spans between the initial and final nodes of the segment. The length of the segment is defined as the integrated arc length of its trajectory between these two nodes. The diameter of the segment is in general a function of position along the trajectory. Two simple models are to assume (1) constant diameter, or (2) a tapering diameter characterized by

the initial and final diameters.

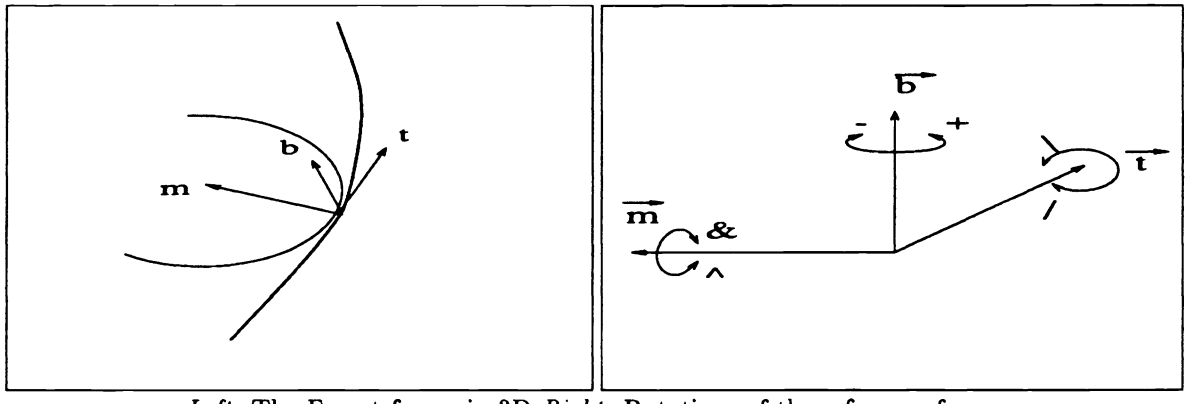
3.2.2 The segment trajectory defines a local reference frame

Segments may be visualized as a tube with a fibrous internal structure (e.g., microtubules) which, like a plant stem, may impart stiffness against twisting. To accommodate this contingency we define a local coordinate system, or frame, with origin at any point **P** along its trajectory. The first axis is selected to lie along the unit *tangent vector* **t** at a point **P**; the other two axes lie in the cross-sectional plane at **P**, whose normal is **t**. The second axis with unit vector **b**, called the *binormal vector*, is chosen in this cross-sectional plane to reflect the orientation of the segment, or tube, about its tangent vector. (We can think of the **b** vector, as it sweeps along the trajectory, as "painting" a longitudinal stripe along the tube where it pierces the tube.) The third orthonormal vector of the right-handed coordinate system, **m**, called the *main normal vector*, is given by $\mathbf{m} = \mathbf{b} \wedge \mathbf{t}$, where "$\wedge$" represents a cross product.

But for dendritic or axonal segments, in the absence of a longitudinal stripe or access to the internal distribution of microtubules, how does one define a local frame along the segment trajectory in some natural way from the available information? The answer is to evoke a special local cartesian coordinate system, called the *Frenet frame*, linked to a point $\mathbf{x}(s)$ on segment trajectory, as parameterized by s.[9] The Frenet frame is defined by the triplet (**t m b**) of orthonormal vectors, where we retain the same names for these vectors as above; in particular **t** is the tangent vector (Figure 2 *left*). These vectors are now computed from the intrinsic geometry of the trajectory:

$$\mathbf{t} = \frac{\dot{\mathbf{x}}}{||\dot{\mathbf{x}}||}, \quad \mathbf{m} = \mathbf{b} \wedge \mathbf{t}, \quad \mathbf{b} = \frac{\dot{\mathbf{x}} \wedge \ddot{\mathbf{x}}}{||\dot{\mathbf{x}} \wedge \ddot{\mathbf{x}}||} \tag{1}$$

where ".": means derivative with respect to s, and all required derivatives are assumed to exist. In essence the binormal **b**, computed from the local curvature, serves to define the "longitudinal stripe" of our illustration. The geometric interpretation of these vectors is straightforward. The vector **t** points along the trajectory at $\mathbf{x}(s)$; **b** is the normal to the plane containing the osculating circle which passes through the point $\mathbf{x}(s)$; and **m** points to the center of the osculating circle. For curves traditionally used in computer graphics, such as Bezier curves, the Frenet frame is readily computable, as also are the curvature κ and torsion τ at $\mathbf{x}(s)$.



Left: The Frenet frame in 3D *Right:* Rotations of the reference frame.

Figure 2: The reference frame

In summary the reference frame is defined along the segment trajectory. Its computation requires specification and parameterization of the segment trajectory but makes no reference to the additional information necessary to specify the generalized cylinder of the segment.

3.2.3 Segment placement in the dendritic or axonal arbor

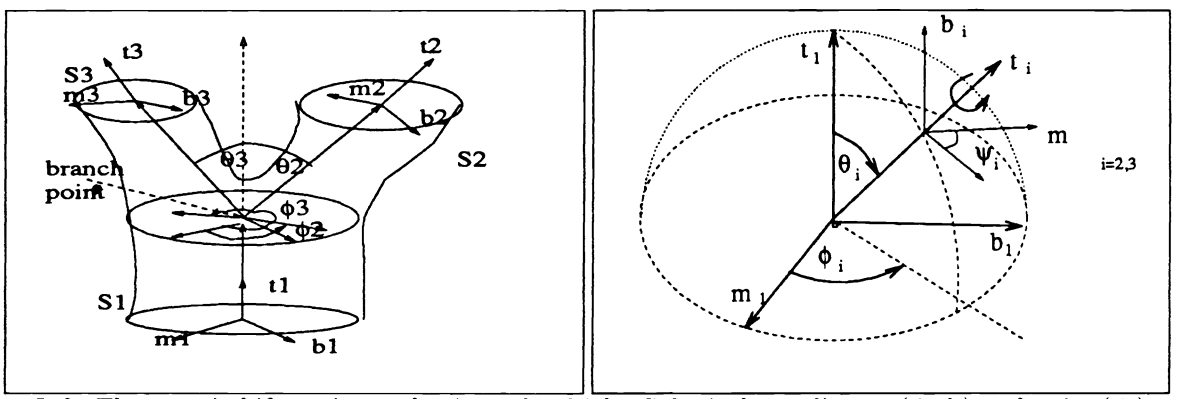
The dendritic (axonal tree) is fully characterized by its segments. A segment, in turn, is represented by (1) its *trajectory*, (2) the additional information to specify its *generalized cylinder*, and finally (3) its *placement* in the dendritic (axonal) arbor. For the trajectory this information consists of the initial and final nodes, the trajectory between these nodes and its arc length, the initial and final reference frames, and the integrated torsion, or twist, of the local reference frame as it traverses the trajectory. The additional information to specify the generalized cylinder consists of the diameter, which may taper along the segment, and the distribution of spines and bouton on the surface of the generalized cylinder, if observed. Finally tree placement information consists of segment

order (stem or first, second, ...), and whether the segment is terminal or non-terminal, as characterized by the final node of the segment.

3.3 Representation of junctions

3.3.1 Geometry of the bifurcation model

A generic bifurcation is exhibited in Figure 3 *left*. The trajectories of the parent segment $\mathcal{S}_1$ and the two daughter segments, $\mathcal{S}_2$ and $\mathcal{S}_3$, meet at a common point called the branch point of the junction. In Figure 3 *left* these segments are shown truncated. The length and diameter of these stubs, however, is irrelevant, as all significant properties of the junction are extrapolated to the branch point.* The order of a junction is defined to be that of its parent segment $\mathcal{S}_1$.



Left: The generic bifurcation and exit angles *Right:* Spherical coordinates (ϕ_i,θ_i) and twist (ψ_i) describing relative orientation of the reference frame for segment $\mathcal{S}_i$ (i = 2, 3)

Figure 3: The bifurcation in spherical coordinates

Conceptually the bifurcation can be modeled as a physical piece of plumbing, a "Y"-joint, into which three segment "pipes" $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ can be screwed. By screwing in the pipes and tightening, each pipe is given a terminal orientation about its axis, defined by the binormal vector $\mathbf{b}_i$ (i=1,2,3) of the segment at the branch point. Apart from these three terminal segment twists, we can give an invariant geometric description of the junction as a rigid body by specifying the angles θ_{12}, θ_{13}, and θ_{23} between the three segment tangent vectors and its chirility C (-1, 0, 1). Bifurcations with opposite chirility are mirror reflections of one another. Specifically, if the segments $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ are modeled respectively by the wrist, thumb and first finger of the right hand, then the chirility is 1; if $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ lie in a common plane, the chirility is 0;, and if left-handed, -1 (Figure 4).

It proves convenient to express the tangent vectors $\mathbf{t}_2$ and $\mathbf{t}_3$ of the daughter segments in spherical coordinates (θ_2,ϕ_2) and (θ_3,ϕ_3) respectively in the reference frame defined by the parent segment, with polar axis $\mathbf{t}_1$(Figure 3 *right*):

$$\mathbf{t}_1 = (0,0,1) \tag{2}$$

$$\mathbf{t}_i = (\sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i) \text{ for i=2, 3} \tag{3}$$

Hence we derive

$$\cos\theta_{23} = \mathbf{t}_2.\mathbf{t}_3 = \cos\theta_2 \cos\theta_3 + \sin\theta_2 \sin\theta_3 \cos(\phi_2 - \phi_3) \tag{4}$$

and observe

$$C = -\text{sgn}(\mathbf{t}_1.(\mathbf{t}_2 \wedge \mathbf{t}_3)) = \sin\theta_2 \sin\theta_3 \sin(\phi_2 - \phi_3) \tag{5}$$

is uniquely defined by the invariant geometry of the junction.

Recapping, the bifurcation is fully specified by two 3-dimensional rotations, $\mathcal{R}_2$ and $\mathcal{R}_3$, about the branch point which transform the final reference frame of parent segment $\mathcal{S}_1$ into the initial reference frame of daughter segment $\mathcal{S}_2$ and $\mathcal{S}_3$ respectively. In particular, these rotations specify how the tangent vector $\mathbf{t}_1$ of the parent segment at the branch point is transformed in the initial segment tangents $\mathbf{t}_2$ and $\mathbf{t}_3$ respectively. Segments, however, also have thickness and internal structure (e.g., microtubular structure) which may define an intrinsic

*Mathematically speaking, the three trajectories characterizing $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ could fail to extrapolate to a common branch point. We ignore this source of error as being insignificant in practice

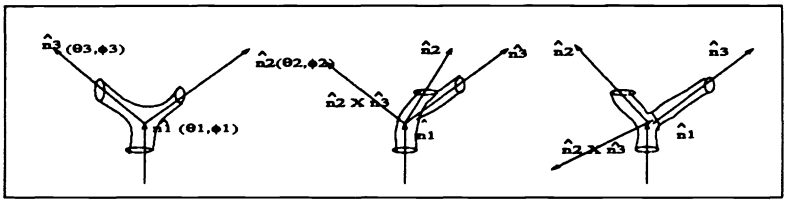


Figure 4: Chirility

orientation of the segment about its tangent. Accordingly the rotation $\mathcal{R}_i$ (i = 2, 3) specifies (1) the rotation of $\mathbf{t}_1$ into $\mathbf{t}_i$, and (2) the final 2-D rotation (twist ψ_i) about $\mathbf{t}_i$ which sets up the initial alignment of the daughter segment $\mathcal{S}_i$ (Figure 3 *right*).

Formally, let the rotation $\mathcal{R}_i(\phi_i, \theta_i, \psi_i)$ about the branch point transform the reference frame $F_1(\mathbf{t}_1, \mathbf{m}_1, \mathbf{b}_1)$ of the parent segment into the reference frame $F_i(\mathbf{t}_i, \mathbf{m}_i, \mathbf{b}_i)$ (i = 2, 3) of daughter segment $\mathcal{S}_i$. The rotation is expressed as the product of three successive 2-D rotations:

$$\mathcal{R}_i(\phi_i, \theta_i, \psi_i) = \mathcal{R}_t(\psi_i)\mathcal{R}_b(\theta_i)\mathcal{R}_t(\phi_i) \tag{6}$$

where ϕ is the azimuthal rotation about $\mathbf{t}_1$, and θ is the latitudinal rotation about $(\mathbf{t}_1 \wedge \mathbf{t}_i)/\|\mathbf{t}_1 \wedge \mathbf{t}_i\|$, the rotated $\mathbf{b}_1$ binormal. These two rotations carry $\mathbf{t}_1$ into $\mathbf{t}_i$. Finally ψ_i is the rotation (twist) about the $\mathbf{t}_i$ axis which defines the initial orientation of segment $\mathcal{S}_i$. If $\mathcal{S}_i$ has a straight or piecewise linear trajectory this final "twist" about $\mathbf{t}_i$ is not needed, and $\psi_i = 0$.

3.3.2 Extension to multifurcations

Multifurcations, those junctions in which more than two daughter segments participate, are relatively rare. For example, almost all junctions in motoneurons of the rat are bifurcations, while trifurcations account for 1%.[7]

We observed above that the rotation $\mathcal{R}_i(\phi_i, \theta_i, \psi_i)$ to rotate from the parent reference frame $\mathcal{F}_1$ to $\mathcal{F}_i$, the reference frame of the daughter segment $\mathcal{S}_i$, could be described as the product of three successive axial rotations. We retain the same convention for multifurcations, though the number of daughter segments is now greater than two.

3.3.3 Diameter dependence and Rall's ratio

By convention we designate segment $\mathcal{S}_2$ to be the daughter segment of largest diameter. Thus $\mathcal{S}_2$ serves as the principal limb emerging from the parent segment at the junction.

Rall has predicted on the basis of a passive electrical model of dendrite segments that R, called Rall's ratio, given by $(\Sigma(\mathcal{D}^{3/2}_{daughter}))/(\mathcal{D}^{3/2}_{parent})$ should be 1. For motoneurons R is approximately one.[10] We assume a default value of R = 1 for Rall's ratio in the absence of direct evidence from manually-traced neurons

3.4 Representation of somata and their environment

For synthetic neurons, similar in scale to those shown in section 6, the soma can be modeled by an ellipsoid or a sphere, or for pyramidal cells by a truncated cone. These axially symmetric models require two form factors[7]:

- *soma cross-sectional area* ($\mathcal{SCA}_s$) - from 2D projection onto the sagittal plane, and
- *soma diameter* ($\mathcal{D}_s$) - the diameter of a circle with the same cross sectional area.

The soma, as a rigid body, requires a soma position and orientation relative to the rostral (R), lateral (L), and dorsal (D) anatomical axes of the nucleus (Figure 5), or for cortical neurons, relative to the local cortical surface normal.

At higher resolution, and for the bulbous somata of the autonomic nervous system, we suggest modeling somata by conventional 3D reconstruction. Somata, of course, can be viewed as particularly plump segments and like segments, represented as generalized cylinders about a central trajectory. We suggest that this trajectory initiate at the axon hillock and pass through the centrum of the soma. However, so few soma have been reconstructed in 3D that we are unable to give further guidance at this time.

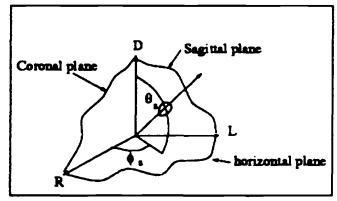


Figure 5: Position and relative orientation of soma

4 OVERVIEW OF L-SYSTEM MODELING

4.1 L-system grammar for generate strings of terminal symbols

A grammar consists of a partitioned set of symbols (*non-terminal* symbols including the *start* symbol and *terminal* symbols) and a set of replacement rules called *productions*. In the context-free grammars considered here, the left hand side of a production designates a non-terminal symbol: its right hand side designates a finite string of symbols, optionally both terminal and non-terminal. To apply a production in an L-grammar, all non-terminal symbols in the string generated so far which match the non-terminal symbol of the left hand side of the production are replaced *in parallel* by the right-hand-side of the production. Paired brackets, as terminal symbols, may be embedded in the string recursively. Productions are applied until the string so generated contains only terminal symbols.

It is convenient to map a string generated by the L-system into an equivalent tree format. A bracketed list of terminal symbols, where the bracketed are paired, can be mapped into a finite tree as follows: the root of the tree is labeled by the start symbol, each non-terminal node of the tree corresponds to a (bracketed) list of terminal symbols, and finally terminal nodes are labeled by terminal symbols of the string while preserving their order in the string

4.2 Turtle rotations are defined in the local reference frame

The terminal symbols of the L-grammar are interpreted graphically as commands for turtle reorientation and movement. These symbols instruct the turtle to trace a path in the 3-dimensional space. The term "turtle" is traditionally used as the turtle conveys not only a position in space, but also its orientation; that is, the turtle defines a local frame of reference ($\mathbf{t}$, $\mathbf{m}$, $\mathbf{b}$) which we shall identify with the reference frame.

Six terminal symbols are reserved to signify rotations of the turtle in its local reference frame (Figure 2 *right*). These are ("\", "/") for tangential rotations about $\mathbf{t}$; ("&", "^") for rotations about main normal vector $\mathbf{m}$; and ("+", " -") for rotations about the binormal $\mathbf{b}$.

4.3 The reference frame can be generalized to piecewise linear trajectories

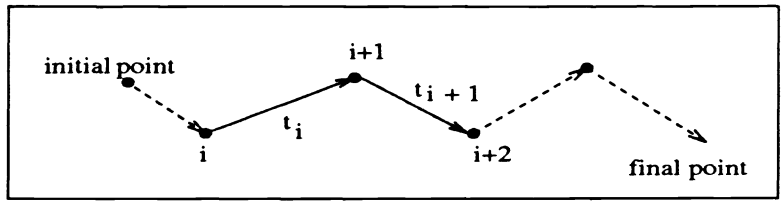


Figure 6: Piecewise linear trajectory

Piecewise linear trajectories occur as approximations to segment trajectories and as describing the path between two nodes in the dendritic or axonal tree. For piecewise linear segment trajectories (Figure 6), the initial and final points of the trajectory are called external knots: all other knots of the piecewise linear curve are called internal knots.The tangent vector $\mathbf{t}$ is defined as a piecewise constant, which changes discontinuously at internal knots. But $\mathbf{b}$, and hence $\mathbf{m}$, are not computable from Eqn 1. Therefore we introduce the following generalization of the reference frame for piecewise linear trajectories. At an internal knot define $\mathbf{b}$ and $\mathbf{m}$ as follows

$$\mathbf{b}_i = \frac{\mathbf{t}_{i-1} \wedge \mathbf{t}_i}{||\mathbf{t}_{i-1} \wedge \mathbf{t}_i||}, \quad \mathbf{m}_i = \mathbf{b}_i \wedge \mathbf{t}_i \tag{7}$$

Then define the generalized reference frame along the trajectory as a piecewise constant function: the reference frame from the initial point is translated without rotation to the second knot point, at this knot **t**, **m** and **b** are computed by Eqn 7 and this reference frame propagated to the next knot, and so on. The final node carries the reference frame computed at its next-to-last knot. For straight trajectories and their generalization, piecewise linear trajectories, the initial orientation of the binormal vector **b** cannot be computed from the segment trajectory alone but requires reference to the branch point and parent segment from which the daughter segment arose. We shall define **b** to be normal to the plane defined by these two segments. Finally for a stem segment having a piecewise linear trajectory, we assign an initial orientation to the **b** vector derived from the coordinate system used to describe the soma.

4.4 L-systems can generate geometric models

An L-system grammar generates a parenthesized strings of terminal symbols. Each terminal symbol serves to orient or generate a geometric model of a constituent part of the organism. For modeling dendritic or axonal trees, the principal constituent parts are segments, which are realized as generalized cylinders and may be of several classes, and junctions, from which grow subtrees. Figure 7 illustrates the generation a a simple junction

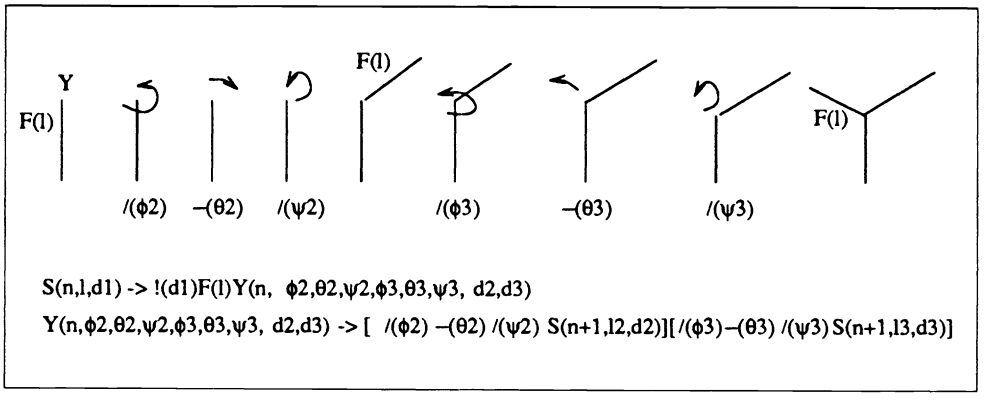


Figure 7: A simple grammar that generates a junction

by an L-system. In this simple grammar, the segments are straight black lines.

4.5 Depth-first generation of the dendritic tree

Terminal symbols which represent segments are realized as generalized cylinders swept out as the turtle traverses the trajectory of the segment.

At terminal symbols which represent junctions (signalled by a "[" symbol in the string representation; and equivalently by a non-terminal node in its tree representation) the turtle is reoriented as specified by the first of the list of rotations specifying the junction. The daughter segment $\mathcal{S}_2$, and its associated subtrees, are swept out in a depth-first traversal of the tree. The turtle, backtracking at each occurrence of a "]" symbol, finally returns to its prior orientation at the original junction, the reference frame of segment $\mathcal{S}_1$. Then the turtle begins anew to traverse in a similar manner the subtrees blossoming from $\mathcal{S}_3$ (and other segments, if a multifurcation). In this manner the turtle completes a depth-first traversal of the tree and by this process generates a geometric model of the neuron.

4.6 Stochastic L-systems generate populations of neurons

A stochastic L-system proceeds as above, but the value of segment parameters and the rotation angles emerging from a branch point are now random variables drawn from appropriate distribution function. In short the terminal symbols designate the appropriate distribution function, as for example the length of a given kind of segment.

In general these distribution functions will depend not only on neuron type, and possibly its the depth in the cortex, but also on the type of dendritic arbor. For example apical and basal dendritic trees of a pyramidal cell are generated by different subgrammars, and will have different distribution functions controlling segment length and branch point exit angles.

5 STOCHASTIC GENERATION OF NEURONS

5.1 Premises

Stochastic generation of neurons makes several simplifying hypotheses:

- Each dendritic tree develops independently.
- Segments develop independently, though they may exhibit tropism in the presence of a shared external force field or conform to a common nucleus boundary.
- Segment length is a random variable whose distribution function is both class and order dependent.
- Each class of segments of order n is assigned a probability of non-termination $\mathcal{P}_{nt}$(n). Typically $\mathcal{P}_{nt}$(n) is a monotonically decreasing function of n, as illustrated in Figure 8.

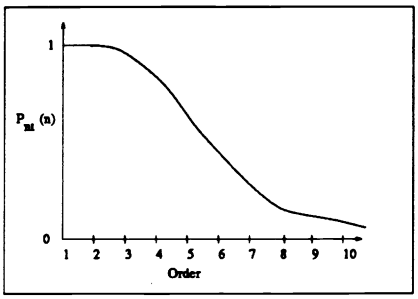


Figure 8: Typical probability that segment is non-terminal vs segment order

- Bifurcations and multifurcations develop independently, but are conditional upon the class of the parent and daughter segments.
- The junction probability density function describing the distribution of branch point exit angles (and twists, if any) can order dependent.
- Both segment and junction distribution functions can depend on the position of the soma in the neural tissue, (e.g., pyramidal cells exhibit scale dependence upon their depth in the cortex).

5.2 Generation of segments

The stochastic generation of segments attempts to simulate the growth process of the neurobiological segment but from a coarser grain perspective. We distinguish three models of the environment for this growth:

- *the isotropic model*, where in the absence of environmental forces we assume that segment growth is determined solely by its local reference frame;
- *the force field model*, where an external force field provides a tropism for growth. The apical dendritic trees of pyramidal cells, for example, exhibit tropism and grow predominantly toward the cortical surface. Using the context-sensitive, parametric L-grammars, we can also generate neurons whose growth conforms to the surface boundary of an encompassing nucleus. The nucleus boundaries can be specified as implicit functions in the model space. The distance from the surface and the angle made with the surface can be used as parameters to reorient the direction of growth. These parameters are adjusted with each generation of rewriting, and so this mechanism provides a feedback to the L-system about whether or not a reorientation in needed for the next stage of growth.[2]
- *the 3D grid model*, where a boundary-conforming grid is imposed throughout the body of the brain nucleus or architectonic area. The grid imposes, by interpolation, a curvilinear coordinate system throughout the neural tissue. Segment generation is then prejudiced to conform to this local coordinate system. For example, Purkinje cells develop as if conforming to planar glial trellises.

The force field model and the 3D grid model can be viewed as alternative statements of the same model. We sketch this correspondence briefly below. The force field $f(\mathbf{x})$ is assumed acyclic, that is curl $f(\mathbf{x}) = 0$. The force f can then be derived as the gradient of a scalar potential Φ, that is $f = \nabla\Phi$. Now we design the potential Φ such that the nucleus boundary is an iso-surface of Φ, and hence $\nabla\Phi$ is the local inward-pointing surface normal. Setting $f = \nabla\Phi$ sets up a repulsive force field, which bends ramifying segments of the dendritic tree to conform to the nucleus surface (e.g., the cortical surface).

5.3 Generation of junctions

An individual junction has been represented, as described above, as a list of rotations, where each rotation in turn transforms the reference frame of the parent segment into the reference frame of a daughter segment. The stochastic generation of junctions (of given order and segment classes) requires that we specify the conditional probability density for these concurrent frame rotations.

To illustrate the concept, consider a bifurcation for which the incident trajectory bifurcates ("scatters") into daughter trajectories $(\phi_2, \theta_2, \psi_2)$ for $\mathcal{S}_2$ and $(\phi_3, \theta_3, \psi_3)$ for $\mathcal{S}_3$. The differential probability for this event is

$$\mathcal{P}(\phi_2, \theta_2, \psi_2, \phi_3, \theta_3, \psi_3 | n, \mathcal{S}_2, \mathcal{S}_3) \mathrm{d}\omega_2 \mathrm{d}\psi_2 \mathrm{d}\omega_3 \mathrm{d}\psi_3 \tag{8}$$

where $\mathrm{d}\omega_2$ and $\mathrm{d}\omega_3$ are the differential solid angles given by

$$\mathrm{d}\omega_i = \sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i \text{ for i=2, 3} \tag{9}$$

The observant reader will recognize that the statistical distribution of junctions, as necessary for their stochastic generation, directly parallels the S-matrix theory for particle scattering in theoretical nuclear physics.[11] In fact, if segments have an inherent orientation (as defined by their reference frame) then this additional complication plays a role comparable to polarization for particle beams (which also require the definition of a local reference frame for each beam). However we shall defer further development of this correspondence between the junction distribution functions and the S-matrix formalism until subsequent papers.

5.4 Generation of a forest

Typically only a sparse "forest" consisting of 0.01-0.1% of the cells in a nucleus/architectonic area (in the rat) are displayed simultaneously, with the choice of cells to be displayed under investigator control. We offer different classes of non-uniform distributions of neurons in 3D space, such as the Poisson, the Poisson sphere and the jittered distributions. A 'seed' for a neuron may be planted in 3 D space according to the distribution. This allows an investigator to conduct studies with controlled variation in population densities of neurons.

An additional and particularly significant environmental influence must be built into the L-grammars when parallel growth of neurons is being considered. An important inviolable constraint is imposed by the fact that space is not a shared resource. So one must prevent processes of one neuron from encroaching into space occupied by some other during growth. We call this *the intersection problem.* Intersections can be avoided by building in feedback in the L-grammar about the current state of the string generated. Each production must be applied conditionally, constrained that the result of the production will not result in the violation of the space rights. We are making the necessary modifications in the *lsys* software for this purpose.

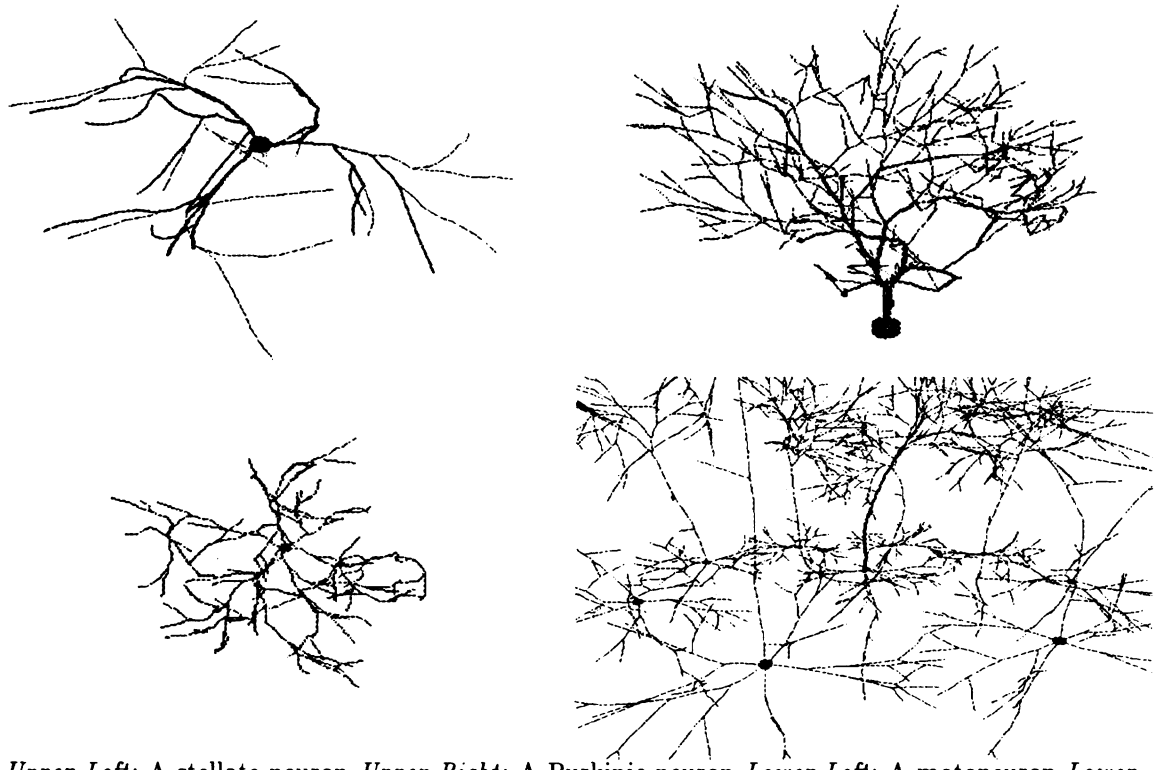
6 GROWTH OF SAMPLE NEURON POPULATIONS

6.1 Introduction

The L-grammar we use consists of terminal symbols that are commands for the turtle. The turtle can turn around its three principal axes and can thus point into the direction in which the primary and secondary segments need to be grown. The growth of the primary and secondary segments at the branch point involves calculation of the exit angles by which the turtle needs to turn, pitch and then twist before sending off the turtle in thiis new orientation. Since neurons are not like chemical molecules, these angles are distribution functions rather than single fixed values (delta function distributions). These angular distributions must be provided by the statistical analysis of the measured neuron morphology databases. For instance, through a statistical analyzer it may be discovered that the branching angle between the primary and secondary segments is a normal distribution with a mean of 60 degrees and a standard deviation of 20 degrees for a given neuron type. Or it may be found that the primary and secondary segments have absolutely no idea of each others growth and bifurcate quite independently from the branch point. Mathematically speaking, their exit angles are uncorrelated, and their probability density function, $\mathcal{P}(\phi_2, \theta_2, \psi_2, \phi_3, \theta_3, \psi_3) = \mathcal{P}_2(\phi_2, \theta_2, \psi_2), \mathcal{P}_3(\phi_3, \theta_3, \psi_3)$, factors into independent probabilities of each daughter segment. These conclusions can be used in the stochastic L-grammars to provide values for turn, pitch and twist of the turtle.

6.2 Pyramidal, motoneuron, and Purkinje cells

Pictures of different neuron types that were generated using the stochastic L-system synthesizer after estimates of the distribution functions were supplied, are shown in Figure 9.



Upper Left: A stellate neuron *Upper Right:* A Purkinje neuron *Lower Left:* A motoneuron *Lower Right:* A Pyramidal forest

Figure 9: Sample neurons

The current grammars assume normal distribution functions. The mean and standard deviation of the parameters are supplied through the grammar. For instance, if the the distribution of ϕ_2 , θ_2 and θ_3 are known, then the angle θ_{23} can be computed. However as the angles are extracted from a distribution their values have to be checked against geometrical constraints like $\theta_{12} + \theta_{13} + \theta_{23} \leq 360$ and $\theta_{23} \geq \text{abs}(\theta_2 - \theta_3)$.

The diameters of the daughter segments and the parent segment are calculated from the parent segment's diameter, the distribution estimate for the diameter of the primary segment and Rall's ratio set to 1.

A Purkinje neuron has a relatively planar architecture. Even though its individual segments are not in a perfect plane, the cell somehow maintains a mean planar character as if there were two invisible walls between which this cell is sandwiched. To control the individual twists from destroying this overall planar character and yet avoid perfectly planar dendritic arborization, the turtle is influenced by a kind of tropic effect. The current turtle direction is mirrored with respect to the normal of the plane of the Purkinje cell. This forces the segments of the alternate orders to return back to the plane and hence prevents a propagating deviation from destroying the planarity of the cell.

A typical rat motoneuron has from 6 to 12 first order dendritic segments, with the mean diameter ranging from 4.4 to 6.8μm.[7] The maximum segment order ranges from 8 to 12. Almost all branching is dichotomous. On an average, each dendrite has 27 segments, 13 branch points, and 14 terminations. By calculating the ratio of the number of segments at order n to that at order n+1, one can estimate the probability of non-termination. This estimate was used in the stochastic application of productions of the L-grammar to produce the motoneuron shown in Figure 9.

Once a grammar model for a particular neuron morphology is perfected, it can be used to generate a population of its types. Since the parameters of the grammar are pulled out from their distribution functions, no two cells will be identical yet they will all be similar. Figure 9 shows such a population of pyramidal neurons.

As an application of the force field model, consider the flat nature of the apical dendritic arbor of a pyramidal neuron morphology. To achieve this flat character, a vector field along the vertical axis of the cell was applied at the point when the turtle is generating segments of this flat arbor. If the turtle is made to reorient itself such that it is always headed in the direction perpendicular to the vector field, the required flatness is simulated. This realignment force is an exponential function of the turtle's position from the surface; the effect being negligible away from the surface and strongest at it.

Currently only spherical boundaries are available, but other shapes can be easily implemented. These shapes are also specifiable in the grammar rules and can be propagated or bracketed in order to achieve global as well as local boundaries.

6.3 Stochastic seeding of neuron populations

The population morphological data can be obtained in two ways:

1. by invoking the same grammar as many times as the number of required cells; or
2. by introducing a higher level production in the grammar which distributes the neurons by walking the turtle in model space and planting the somata of the neurons. This gives a stochastic L-system grammar for a neuron population.

Although the latter approach seems more elegant, the system and memory restrictions may force the first approach to be adopted if the number of neurons required is fairly large.

Different seeding distributions can be invoked from within the grammar by adding these distribution functions to our neuron generator, just like the normal distribution has been made a part of the system. Some such distributions are the Poisson, the Poisson Sphere and the jittered distributions. In case the second approach is used, these distribution functions can be used to walk the turtle to generate the population.

7 NEURON MORPHOLOGY MODELER

Experiments with various stochastic L-system models of neuron morphology have given us neurons with promising proximity in shape to neurons illustrated in the neurobiology literature. Work is in progress to improve this degree of proximity, but more importantly to validate the L-system neuron models against databases of manually-traced neurons. To this end the Neuron Morphology Modeler (NMM), described below, has been devised to provide for iterative refinement of L-system models (Figure 10).

7.1 Data level of the modeler

Central to the scheme at the data level is an *object-oriented neuron morphology database* (Paradox, Borland International, Scotts Valley, CA). This DBMS stores for each neuron a representation of its soma and for each dendritic or axonal tree a representation of its segments and junctions, as described above.

This DBMS can receive input (at the data level) from the *neuron tracer* (Neuron Tracing System, Eutectic Electronics, Inc., Raleigh, NC) or (at the knowledge level) from the L-system *neuron generator* (a derivative of public-domain software known as *lsys*, copyright (c) 1990, 1991 by Jonathan Leech, University of North Carolina). In this sense the DBMS is indifferent as to whether it is storing data from real manually-traced neurons or L-system-generated synthetic neurons.

Upon command, the DBMS can prepare database files for *the spreadsheet* program (Excel, Version 4.0, Microsoft Corporation). The spreadsheet program allows data editing and graphical summarization.

Spreadsheets, in turn, providethe conventional input to the *statistical analyzer* (Systat, Version 5.03, Systat Inc., Evanston, IL). In particular the statistical analyzer estimates the distribution functions for segment non-termination, segment length and trajectory generation, and the rotations characterizing bifurcations. At present only ad hoc techniques are available for using Systat for this purpose.

7.2 Knowledge level of the modeler

We now enter the knowledge level (Figure 10) where the parametric statistical dependencies and distribution functions ascertained by the statistical analyzer are used to refine the L-system models of the neuron population. The *L-system synthesizer* is presently little more than a convenient name for the collection of ad hoc rules which effect the L-system modifications.

Next, paralleling the neuron morphology DBMS, we have at the knowledge level the *neuron morphology knowledge base*, which stores the L-system grammars and their associated distribution functions for stochastic

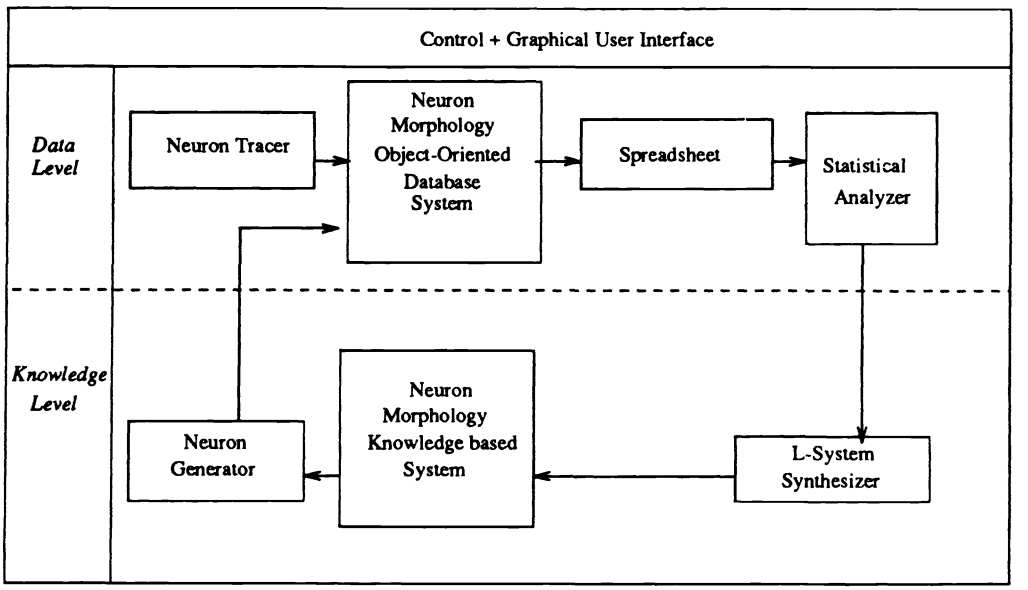


Figure 10: Neuron morphology modeler

generation of synthetic neurons.

And finally, closing the loop, we return to the *neuron generator*, described above.

With experience we anticipate that the system components at the knowledge level will merge into an expert-system-driven simulation system. The data level NMM provides the external environment for this knowledge-level subsystem. The long term goal, of course, is for the knowledge level subsystem to converge upon a stochastic L-system model such that the input from real manually-measured neurons and synthesized neurons can not be statistically distinguished.

7.3 Visualization facilities

Display of neurons (manually-traced or synthetically generated) in 3D uses the principal application programming interface in graphics (GL, Silicon Graphics, Inc.) running on either an Iris or Onyx graphical workstation. The neuron generator *lsys* outputs segments in the turtle order to an output file. This file, after filtering, drives GL. We anticipate moving the visualization facilities to OpenGL, which will be available on the PC. Our goal is to run the entire neuron morphology modeler on a single machine, a high-end PC.

Additional visualization software developed by Brent Burton, Texas A&M University, changes the stick figures of *lsys* into 3D volumes. Burton is also developing a navigator to traverse graphical models of brain nuclei and other neural tissue as viewed in animation. This visualization facility requires the Onyx workstation.

8 SIGNIFICANCE

A representational framework for neuron morphology is presented that is adequate both for the quantitative description and stochastic generation of neuron populations. All models are presently limited to modeling neuron morphology at the limit of optical resolution.

The applications of these stochastic L-system models are several. The L-system models bring coherence to existing neuron databases; they allow us to estimate the new information contributed by the additional neuron morphological measurements.

A paradigm shift to a new mode of analyzing neuron morphology is suggested. The developing bifurcations and multifurcations of growing dendritic and axonal arbors bear striking resemblance, in ultra-slow motion, to nuclear scattering events. We recognize this resemblance, and statistically describe these junction "events" by a variant of the S-matrix theory used in the formal theory of nuclear reactions. Conceptually, the block of embedded histological tissue replaces the bubble chamber of high energy physics. The methodology is now largely in place for the measurement and morphological analysis of large neuron populations.

9 REFERENCES

[1] Rozenberg G., Salomaa A. (Eds), "Lindenmayer Systems", Springer-Verlag, 1992.

[2] Prusinkiewicz P., Lindenmayer A., "The Algorithmic Beauty of Plants", Springer-Verlag, New York, 1990.

[3] Kreowski Hans-Jörg, "Parallel Hyperedge Replacement", "Lindenmayer Systems", Rozenberg G., Salomaa A.(Eds.), Springer-Verlag, Berlin, 1992.

[4] Lindenmayer A., Jürgensen, "Grammars of Development: Discrete-State Models for Growth, Differentiation, and Gene Expression in Modular Organisms", "Lindenmayer Systems", Springer-Verlag, New York, 1992.

[5] Böers E., Kuiper H., "Biological Metaphors and the Design of Modular Artificial Neural Networks", Master's thesis, Department of Computer Science and Experimental and Theoretical Psychology at Leiden University, Netherlands, 1992.

[6] Lindenmayer A., "Mathematical Models for Cellular Interaction in Development, Parts I and II.", Journal of Theoretical Biology, 18:280-315, 1968.

[7] Chen X.Y.,Wolpaw J. R., "Triceps Surae Motoneuron Morphology in the Rat: A Quantitative Light Microscopic Study.", J. Comp. Neurol., 343:143-157, 1994.

[8] Ohara P.T., Havton L.A., "Dendritic Architecture of Rat Somatosensory Thalamocortical Projection Neurons", J. Comp. Neurol., 341:159-171, 1994.

[9] Farin G., "Curves and Surfaces for Computer Aided Geometric Design", Academic Press, San Diego, 1993.

[10] Rall, W., "Core Conductor Theory and Cable Properties of Neurons". In E.R. Kandel (ed): The Nervous System: Cell Biology of Neurons. Section 1: Handbook of Physiology, Part 1, Vol. 1. Bethesda: American Physiological Society, pp. 39-97.

[11] Blatt, J.M., Weisskopf, V.F.,"Theoretical Nuclear Physics", John Wiley & Sons, New York, 1952, Chapter X.

[12] Knupp, P., Steinberg, S.,"Fundamentals of Grid Generation", CRC Press, Boca Raton, 1994.