Reset reproduction of two articles published in *Computational Intelligence*, Vol. 8, No. 2 (May 1992): *Prolegomena to any future qualitative physics*, pp. 187–209, and *Epilegomenon*, pp. 326–335. Reprinted July 1994.

Prolegomena to Any Future Qualitative Physics*

Elisha P. Sacks
Department of Computer Science, Princeton University, Princeton, New Jersey 08544

Jon Doyle
Laboratory for Computer Science, Massachusetts Institute of Technology
545 Technology Square, Cambridge, Massachusetts 02139

Abstract

We evaluate the success of the qualitative physics enterprise in automating expert reasoning about physical systems. The field has agreed, in essentials, upon a modeling language for dynamical systems, a representation for behavior, and an analysis method. The modeling language consists of generalized ordinary differential equations containing unspecified constants and monotonic functions; the behavioral representation decomposes the state space described by the equations into discrete cells; and the analysis method traces the transitory response using sign arithmetic and calculus. The field has developed several reasoners based on these choices over some fifteen years. We demonstrate that these reasoners exhibit severe limitations in comparison with experts and can analyze only a handful of simple systems. We trace the limitations to inappropriate assumptions about expert needs and methods. Experts ordinarily seek to determine asymptotic behavior rather than transient response, and use extensive mathematical knowledge and numerical analysis to derive this information. Standard mathematics provides complete qualitative understanding of many systems, including those addressed so far in qualitative physics. Preliminary evidence suggests that expert knowledge and reasoning methods can be automated directly, without restriction to the accepted language, representation and algorithm. We conclude that expert knowledge and methods provide the most promising basis for automating qualitative reasoning about physical systems.

1 Introduction

> To develop mathematics, one must always labor to substitute ideas for calculations. Dirichlet

The qualitative physics enterprise of artificial intelligence seeks to automate reasoning about the physical world in the manner of scientists, engineers, and other experts, ultimately “constructing what could best be described as an ‘artificial engineer’ or ‘artificial scientist’ ” (Weld and de Kleer, 1990, p. 1). Most of this research seeks to capture the ability of experts to predict the behavior of dynamical systems, such as circuits, fluid flows, and mechanisms. Experts reason about dynamical systems by formulating and analyzing differential equations that capture the properties of interest and abstract away irrelevant details. Qualitative physics hypothesizes that experts mostly use extremely general equations and analysis tools. The rationale is that general equations are easier to formulate than specific equations and that general analysis is more rapid and robust than detailed mathematical or

*This paper is a revision and expansion of [12]. Authors listed anti-alphabetically.

numerical analysis. Researchers have developed several approaches based on this hypothesis. Although each approach has its own special characteristics, they share a modeling language for physical systems, a representation for behavior, and an analysis method. In this paper, we assess this shared theory of expert reasoning, which we call SPQR (short for *Simulātio Processūs Qualitātivō Ratiōcinātiōne* or "Simulation of Processes by Qualitative Reasoning").

We believe that qualitative physics has not lived up to its initial promise to automate expert reasoning: not because this task is impossible, but due to fundamental limitations of SPQR. After fifteen years of research, SPQR can analyze successfully only a handful of simple systems, such as a falling point mass or a U-shaped tube containing liquid. It struggles with linear oscillatory systems, such as simple springs, and fails completely on many nonlinear oscillators. Despite numerous attempts to improve performance to the level of expert analysis by extending SPQR, the analyses that appear in current research papers seem, from a mathematical point of view, little better than those of five years ago. Qualitative physics blames this state of affairs on the unforeseen difficulty of identifying and automating the general mathematical principles that underlie expert reasoning.[1] We propose a different explanation for the weakness of SPQR compared with routine expert performance: SPQR embodies inappropriate assumptions about the needs and methods of experts for reasoning about dynamical systems, in that it provides little information about asymptotic behavior and uses little expert knowledge.

Experts need to know the asymptotic behavior of a dynamical system: the stable steady-states, the sets of solutions that converge to each steady-state, and the sensitivity of these properties to perturbations in the equations. This information provides a qualitative understanding of the system and sets the stage for further analysis, such as calculating the transient behavior for specific initial conditions. The most important steady-states are constant and periodic behavior, followed by quasi-periodic and chaotic behavior. The SPQR behavioral representation cannot express any information about steady-states, except for the existence of constant solutions, and the SPQR analysis algorithm only traces the transient behavior of a system for a range of initial conditions. SPQR thus fails to answer experts' questions about asymptotic behavior.

Experts reason about dynamical systems with advanced mathematics, physics, and other knowledge. They formulate, analyze, and revise specific equations until predictions derived from these equations become accurate enough for their needs. They eschew reasoning about general equations because these support few useful inferences. In many cases, they analyze equations by linearizing them over the range of interest and applying the well-developed theory of linear systems. When nonlinearity cannot be ignored, they resort to the mathematical theory of dynamical systems, which provides qualitative descriptions of the possible steady-states, and to numerical software for analyzing specific equations. They use the theory to guide the numerical analysis and to verify the results. SPQR, which embodies little knowledge about linearity, dynamical systems theory, or numerical analysis, cannot reproduce expert understanding.

[1] For example, according to Weld and de Kleer (1990, p. 7), "codifying qualitative knowledge about the physical world has proven to be surprisingly difficult. Just getting the qualitative version of the calculus (used to express qualitative knowledge) right took many years of work by a large number of people."

Concerns with producing causal explanations and with explaining commonsense reasoning may underlie the differences between SPQR and expert knowledge and methods, in that some qualitative physics researchers claim that SPQR constructs causal explanations of the workings of physical devices, and some intend qualitative physics to cover commonsense reasoning as well as expert reasoning. But both these positions are controversial.[2] We will not address either of these issues here in order to focus more clearly on the accepted aims and methods of SPQR. Whatever one's position on causality, the limitations of SPQR reduce its utility for producing causal explanations, since it cannot explain systems that it cannot analyze. Using expert methods might thus aid in producing causal explanations if it proves easier to produce causal paraphrases of possibly noncausal explanations than to enhance SPQR. And whether or not the man on the street infers that "what goes up must come down" by reasoning with a general model, the aeronautical engineer analyzes exact aircraft models by advanced mathematics and by extensive numerics. Indeed, a great part of his or her education involves learning how to augment, replace, or refine commonsense concepts and methods (whatever they may be) with more informed techniques.

The plan of discussion is as follows. In Section 2, we describe SPQR and examine its successes, failures, and recent extensions. We demonstrate that most of the successes are syntactic variants of three simple equations. Despite the extensions, SPQR fails on other simple problems. Moreover, we prove that many generalized differential equations convey no useful information. We conclude that SPQR is far from attaining expert performance. In Section 3, we trace the problem to a mismatch between expert reasoning and SPQR's equations, behavioral representation, and algorithms. In Section 4, we explain why advanced mathematics provides the best available basis for automating expert reasoning about the physical world. We argue that qualitative physics research should focus on modeling and on automating existing mathematics, rather than on inventing analysis tools. In the final section, we summarize our arguments and recommendations.

2 The SPQR methodology

The principal qualitative physics approaches to automating the analysis of dynamical systems include de Kleer and Brown's (1984) confluences, Forbus's (1984) QP theory, and Kuipers's (1986) QSIM. Each of these approaches models dynamical systems with time-varying state variables governed by generalized ordinary differential equations. Some applications of these approaches obtain the equations as input, while others derive them by parsing an input domain model, such as a circuit schematic or a systems dynamics component model. Each approach provides a representation for the equations and algorithms for

[2] On the subject of causality, Weld and de Kleer (1990, p. 611) write, "Causality is by far the most fractious topic in qualitative physics." For example, Iwasaki and Simon (1986a; 1986b) deny that SPQR provides any causal information beyond that provided by the standard causal ordering method of econometrics, and de Kleer and Brown (1986) dispute their conclusions. On the subject of commonsense reasoning, Weld (1990, p. 4) writes "The goal of qualitative physics is to make explicit the unspoken intuitions of experts in the physical sciences. I distinguish qualitative physics from the field of naive physics. Qualitative physics is interested in expert reasoning, not in duplicating the common mistakes of novices." Apparently in contrast, Forbus (1990, p. 11) writes "The goal of qualitative physics is to capture both the commonsense knowledge of the person on the street and the tacit knowledge underlying the quantitative knowledge used by engineers and scientists."

inferring properties of the solutions. We abstract the approaches into a single framework, called SPQR, that captures the essential features of the equations, behavioral representation, and analysis algorithm. Although the superficial details of SPQR most closely resemble QSIM, which we find especially clear and precise, our discussion applies equally to the other approaches. Crawford *et al.* (1990) prove the dynamics module of QP theory equivalent to QSIM by implementing a translator from QP theory to QSIM. We prove the applicability of our arguments to confluences in Section 2.3. Every other approach in the literature closely resembles one of these three.

2.1 SPQR equations, behaviors, and algorithms

SPQR state values are ordered (or partially ordered) sets of numbers and intervals. The most common set of values is the set of sign values $(-\infty, 0)$, 0, and $(0, \infty)$, which we abbreviate as $[-]$, $[0]$, and $[+]$. SPQR state variables map temporal values, which are points or intervals, to SPQR state values. SPQR equations relate state variables via arithmetic operators, differentiation, and functional relations. The arithmetic operators and functions map SPQR state values to SPQR state values. The functions are specified as strictly monotonic increasing (M^+) or decreasing (M^-) and possibly by a few stipulated values. In particular, the symbols M_k^+ and M_k^- denote monotonic functions that take the value 0 at k. Differentiation maps a state variable to its derivative state variable.

SPQR defines the state of a system as the SPQR values of its state variables and of their derivatives. It characterizes the behavior of the system by the sequences of states that it can go through, ordinarily seeking to identify behavioral properties that hold in every state sequence compatible with the equations. It represents state sequences as a transition graph whose nodes and links denote states and possible transitions.

SPQR derives the graph by repeatedly identifying the current state and finding all immediate successor states. It identifies the current state by applying interval arithmetic rules for combining SPQR values, such as $[-] + [0] = [-]$, and propagating these results through the equations. It finds the immediate successors with calculus rules, such as the intermediate value and mean value theorems. For example, the state $x = [0]$ and $\dot{x} = [+]$ may immediately follow $x = [-]$ and $\dot{x} = [+]$, but the state $x = [+]$ and $\dot{x} = [-]$ may not. Full descriptions of typical SPQR algorithms appear in (Kuipers, 1986) and in (Williams, 1984).

2.2 Example: tubular fluid flow

We illustrate the concepts of SPQR with an example popular in qualitative physics: the inertia-free flow of liquid in a U-shaped tube, displayed in Figure 1. We follow the discussion in Kuipers (1986), but rewrite his equations in standard notation. The SPQR description

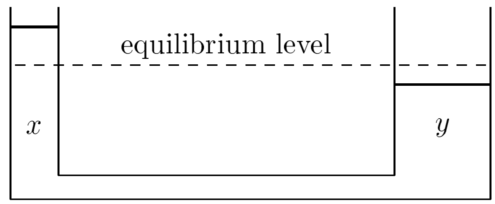

Figure 1: Fluid flow through a U-shaped tube.

of this system consists of the equations

$$\exists e, f, g, h \in M_0^+ \left\{ \begin{array}{rcl} l_x & = & g(v_x) \\ p_x & = & e(l_x) \\ l_y & = & h(v_y) \\ p_y & = & e(l_y) \\ r & = & f(p_x - p_y) \\ \dot{v}_x & = & -r \\ \dot{v}_y & = & r \end{array} \right. \tag{1}$$

with state variables $l_x, p_x, v_x, l_y, p_y, v_y$, and r. The variables l, p, and v measure the levels, pressures, and volumes in the chambers x and y. They take on sign values because they are measured relative to their equilibria.[3] Their values are

$$\begin{array}{c} l_x = p_x = v_x = [+] \\ l_y = p_y = v_y = [-] \end{array} \tag{2}$$

in the state shown in the figure. The variable r measures the flow rate from x to y, and also takes on sign values. Dotted quantities ($\dot{v}_x$ and $\dot{v}_y$) indicate derivatives with respect to time.

SPQR constructs a transition graph for the U-shaped tube as follows. Starting from the initial sign values listed in Equation (2), the analysis algorithm derives that $r = [+]$ from the fifth equation in system (1), since $p_x - p_y = [+] - [-] = [+]$ and $f([+]) = [+]$. It then derives that $\dot{v}_x = [-]$ and $\dot{v}_y = [+]$ from the last two equations. It infers from the properties of derivatives that v_x decreases and that v_y increases. It infers that l_x and p_x decrease and that l_y and p_y increase by applying the chain rule of differentiation to the first four equations. It infers that r decreases analogously. SPQR has now found the entire initial state of the system. It finds a single successor state in which all variables and derivatives equal [0], using the intermediate value theorem and the known relations among variables. It infers that the system cannot leave the [0] state. The transition graph consists solely of these two states (Fig. 2). In physical terms, the fluid moves directly to the equilibrium level and stays there. (It does not oscillate because it has no inertia.)

2.3 SPQR equations versus confluences

De Kleer and Brown (1984) model dynamical systems with sign equations, called *confluences,* such as

$$[\dot{x}] = [x] + [y], \tag{3}$$

[3]Kuipers (1986) uses the true values rather than sign values.

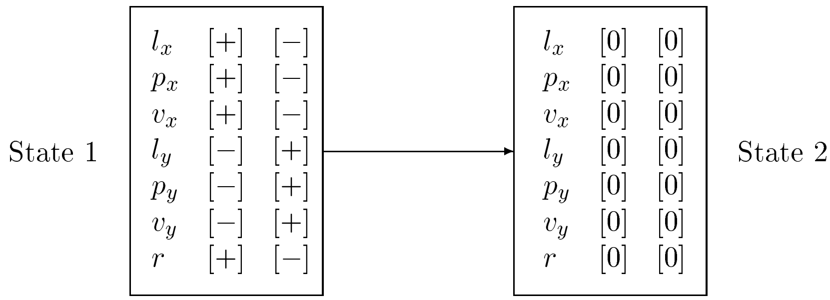

Figure 2: The transition graph of the U-shaped tube. Each variable is listed with its sign and the sign of its time derivative.

rather than with SPQR equations. They interpret confluences as constraints on the signs of the state variable. In equation (3), $\dot{x}$ is positive when one variable is positive and the other is nonnegative; $\dot{x}$ equals zero when both variables equal zero; $\dot{x}$ is negative when one variable is negative and the other is nonpositive; and $\dot{x}$ is unconstrained when one variable is positive and the other is negative. Confluences are more general than SPQR equations because sign expressions represent more functions than do M functions. The sign expression $[x-k]$ with constant k represents every function in $M_k^+(x)$ along with all other functions f satisfying $(x-k)f(x-k) \geq 0$. For example, the function $f(x) = x^3 + 9x^2 + 24x$ is a nonmonotonic instance of $[x]$; it increases from $-\infty$ to a local maximum $f(-4) = -16$, decreases to a local minimum $f(-2) = -20$, increases to $f(0) = 0$, and continues increasing forever. The sign expression $[k-x]$ generalizes $M_k^-(x)$ analogously. Hence, a given confluence subsumes the SPQR equation in which M operators replace sign operators. For example, equation (3) subsumes $\dot{x} = M_0^+(x) + M_0^+(y)$.

We confine attention to generalized equations in the following. The weaknesses of SPQR relevant to SPQR equations arise because these equations are overly general, and so apply all the more to the even more general confluences.

2.4 Extensions to SPQR

Researchers have extended SPQR in many ways. Lee and Kuipers (1988) and Struss (1988) eliminate spurious behaviors that violate the uniqueness of solutions of ordinary differential equations (under the SPQR assumptions). For example, they prove that a block attached to a spring cannot oscillate erratically. Their methods apply only to second-order equations in which time does not appear explicitly. Kuipers and Chiu (1987) and de Kleer and Bobrow (1984) eliminate spurious behaviors based on smoothness assumptions about the higher-order derivatives of state variables. Sacks (1990b) reformulates transition graph construction as proving of algebraic inequalities and develops a less ambiguous algorithm based on an inequality prover and on sign-stability tests. Williams (1988) reduces the ambiguity of analysis algorithms by augmenting SPQR arithmetic with algebraic techniques. For example, he simplifies $x - x$ to 0, whereas sign arithmetic yields the ambiguous expression $[+] - [+]$. Raiman (1986) and Mavrovouniotis and Stephanopoulos (1987) add assertions of the form "variable x is negligible in relation to variable y" and extend the SPQR algorithms to ignore negligible terms. For example, they infer that the sun attracts the earth by neglecting the attraction of the moon. Weld (1990) and Davis (1987) extend SPQR to encompass infinite and infinitesimal values. Kuipers (1987a) forms hierarchies of SPQR equations separated by time-scale. Each level of the hierarchy treats faster levels as instantaneous and treats slower levels as constant. Doyle and Sacks (1991) extend SPQR to predict the relative likelihoods of possible behaviors by viewing the dynamics of a system as a Markov chain over its transition graph. Kuipers and Berleant (1988) derive numerical bounds on the solutions from numerical bounds on the state variables and bounding envelopes around the M functions.

2.5 The state of the art

The qualitative physics literature contains several successful analyses of SPQR equational models of physical systems. The most common examples are fluid flow in a U-shaped tube (Kuipers, 1986), heat flow from a flame to a container of liquid (Forbus, 1990), motion of point masses subject to gravitation and friction (de Kleer, 1977; Kuipers, 1986), and current flow in circuits (Williams, 1984). Kuipers (1985; 1987b) analyzes several homeostatic physiological mechanisms, including the Starling equilibrium that governs the concentration of protein in the body. Molle (1989) analyzes chemical engineering models, such as continuous stirred tank reactors. Falkenhainer and Forbus (1988) analyze a steam plant.

The literature also describes simple physical systems, most commonly a block attached to a spring, whose SPQR analyses are incomplete. The block and spring are modeled by the SPQR equations

$$\exists f, g \in M_0^+ \begin{cases} \dot{x} = v \\ \dot{v} = -f(v) - g(x) \end{cases} \tag{4}$$

with x the displacement of the block from its equilibrium position, v the velocity of the block, f the frictional force, and g the elastic force. SPQR constructs the transition graph shown in Fig. 3. We infer that, after an initial displacement, the block either oscillates around the equilibrium forever or oscillates some number of times and then approaches the equilibrium directly. The transition graph does not specify whether the oscillations die out, remain constant, grow, or vary erratically (Kuipers, 1986). The responsibility for this incompleteness rests with SPQR, since the equations imply that the oscillations die out. We draw this inference by introducing the concept of energy (formalized as a Lyapunov function) and proving that energy decreases to zero along all solutions.

$$\begin{array}{ccccc}
\langle[-],[+]\rangle & \rightarrow & \langle[0],[+]\rangle & \rightarrow & \langle[+],[+]\rangle \\
\uparrow & \searrow & & & \downarrow \\
\langle[-],[0]\rangle & & \langle[0],[0]\rangle & & \langle[+],[0]\rangle \\
\uparrow & & & \nwarrow & \downarrow \\
\langle[-],[-]\rangle & \leftarrow & \langle[0],[-]\rangle & \leftarrow & \langle[+],[-]\rangle
\end{array}$$

Figure 3: Transition graph for the block and spring with $\langle x, v\rangle$ states.

The seeming diversity of successful and partially successful SPQR analyses diminishes upon closer examination. Many elaborate SPQR equations are just syntactic variants of a few simple equations. The simple equations yield the same inferences as the originals with much less work. We find the simple equations by collapsing redundant M functions with the rewrite rules shown in Table 1. The proof that the rules preserve the SPQR semantics is straightforward. Kuipers (1984, App. D) presents an equivalent set of rules.

	expression	**rewrite**		**expression**	**rewrite**
1.	$M^+ + M^+$	M^+	8.	$1/M^-$	M^+
2.	$M^- + M^-$	M^-	9.	$M^+ \circ M^+$	M^+
3.	$-M^+$	M^-	10.	$M^- \circ M^-$	M^+
4.	$-M^-$	M^+	11.	$M^+ \circ M^-$	M^-
5.	$k + M^+$	M^+	12.	$M^- \circ M^+$	M^-
6.	$k + M^-$	M^-	13.	kM^+	M^+ $(k > 0)$
7.	$1/M^+$	M^-	14.	kM^-	M^- $(k > 0)$

Table 1: Rewrite rules for monotonic functions.

To illustrate the use of these rules, we reduce Equation (1), which contains seven equations in four M functions, to one equation in one M function. We obtain $v_x + v_y = k$ by adding the last two equations and integrating. We substitute $k - v_y$ for v_x and combine the remaining equations into

$$\dot{v}_y = f(A(v_y))$$

with

$$A(v_y) = e(g(k - v_y)) - e(h(v_y))$$

by elementary algebra. We rewrite $-v_y$ as $M^-(v_y)$ by rule 3, $k - v_y$ as $M^-(v_y)$ by rule 6, $e(g(k - v_y))$ as $M^-(v_y)$ by rules 9 and 11, $-e(h(v_y))$ as $M^-(v_y)$ by rules 3 and 9, $A(v_y)$ as $M^-(v_y)$ by rule 2, and $f(A(v_y))$ as $M^-(v_y)$ by rule 11. The final result is

$$\dot{v}_y = M^-(v_y).$$

The SPQR analysis of this equation is very easy. In the initial state, $\dot{v}_y = [-]$ because $v_y = [+]$. The only possible successor is $v_y = \dot{v}_y = [0]$. Hence, the volume in arm y increases toward equilibrium. We can recover the values of the other variables from the original equations.

We estimate the number of successful SPQR analyses by surveying the 55 papers in the collection (Weld and de Kleer, 1990), which presents the state of the art through 1989. (We found no new analyses in the proceedings of AAAI-90.) We exclude examples presented without equations, such as the steam plant (Falkenhainer and Forbus, 1988). Table 2 summarizes the results; the full list appears in Appendix A. Out of 37 examples, 30 reduce to three simple equations and 5 are purely algebraic, hence have no dynamics. The U-tube, heat flow, motion of a point mass, and physiological mechanisms are of type 1; the block on a spring and the pressure regulator are of type 2; and the coupled tanks are of type 3. Only two examples are more complicated: an n-Mosfet (Williams, 1984) and a voltage follower (Dague *et al.*, 1987). Although SPQR can analyze types 1, 3, and 4, it cannot fully analyze types 2 and 5. Mathematicians fully analyzed all these equations before the advent of artificial intelligence (Brauer and Nohel, 1969). Hence, only five SPQR equations support useful inferences about dynamical systems and none support new inferences.

type		**#**	**analysis**
1.	$\dot{x} = M^-(x)$	16	full
2.	$\ddot{x} = M^-(\dot{x}) + M^-(x)$	11	partial
3.	coupled pair of type 1	3	full
4.	algebraic	5	full
5.	other	2	partial

Table 2: Number of occurrences of SPQR equations analyzed in Weld and de Kleer (1990).

3 SPQR versus expert reasoning

The methods of SPQR have provided successful analyses of a number of simple systems. But there are other simple and common systems that SPQR still cannot comprehend after 15 years of investigation. From the point of view of classical mathematics, the problems attacked with SPQR today are not in essentials harder, or even different, from those worked on five years ago. Does this simply reflect the preliminary state of research in the field? We believe there is a more fundamental explanation: SPQR embodies inappropriate assumptions about the needs and methods of experts for reasoning about dynamical systems. To support this explanation, we compare the SPQR method of reasoning with the knowledge and methods taught to experts-to-be in standard texts from dynamics (Guckenheimer and Holmes, 1986; Hirsch and Smale, 1974), science (Abraham and Marsden, 1978; Arnold, 1984; Benson, 1982; Clark, 1976; Prigogine, 1980; Thompson and Stewart, 1986), and engineering (Chua *et al.*, 1987; Etkin, 1972; Marsden and Hughes, 1983; Parker and Chua, 1989; White, 1986). We demonstrate that the methods differ profoundly and that the differences explain why experts far outperform SPQR.

3.1 Experts focus on asymptotic behavior

Scientists and engineers normally focus on the asymptotic behavior of dynamical systems. The asymptotic analysis gives a qualitative understanding of the system behavior and sets the stage for further analysis, such as determining the transient behavior and settling time for specific initial conditions. The primary asymptotic concept, called an *attractor,* is a steady-state toward which nearby initial conditions converge. The most important attractors are constant solutions, called *fixed points*, and periodic solutions, called *limit cycles*, followed by *quasi-periodic* solutions that contain two incommensurate periodic components and *chaotic* solutions that wander erratically through state space. Each attractor has an associated *basin of attraction* consisting of the initial conditions that converge to it. The basins of attraction of a system partition its state space into regions of equivalent asymptotic behavior.

Experts reason about asymptotic behavior geometrically and topologically, following the strategy pioneered by Poincaré a century ago. They put differential equations in the normal form

$$\dot{x}_i = f_i(x_1, \ldots, x_n); \quad i = 1, \ldots, n$$

by algebraic manipulation and by introducing new variables as synonyms for higher-order derivatives. For example, the normal form for $\ddot{x} + f(\dot{x}) + g(x) = 0$ is $\dot{x} = v$ and $\dot{v} = -f(v) - g(x)$. They represent the solutions as curves in the Cartesian product of the domains of the state variables, called *trajectories* in the *phase space* of the system. They characterize the (uncountably infinite) solution set of a system by a sketch of its attractors and attractor basins, called a *phase diagram*, thus abstracting away transient behavior and settling times. The qualitative properties of solutions translate into geometric and topological properties of their trajectories.

The block and spring example has a single attractor, the fixed point $x = v = 0$, whose basin is the entire phase space. The phase diagrams of more complicated systems contain other attractors. The trajectories of these attractors are also closed, invariant subsets of phase space. Periodic, quasi-periodic, and chaotic solutions yield simple loops, tori, and fractals. For example, Fig. 4 shows a phase diagram with two attracting limit cycles whose basins are separated by a repelling limit cycle. The corresponding equations model aeroelastic galloping of a square prism in a steady wind (Thompson and Stewart, 1986). The diagram shows that the prism oscillates up and down and that the magnitude of oscillation jumps when the initial condition crosses the unstable limit cycle. Table 3 summarizes these phase space concepts.

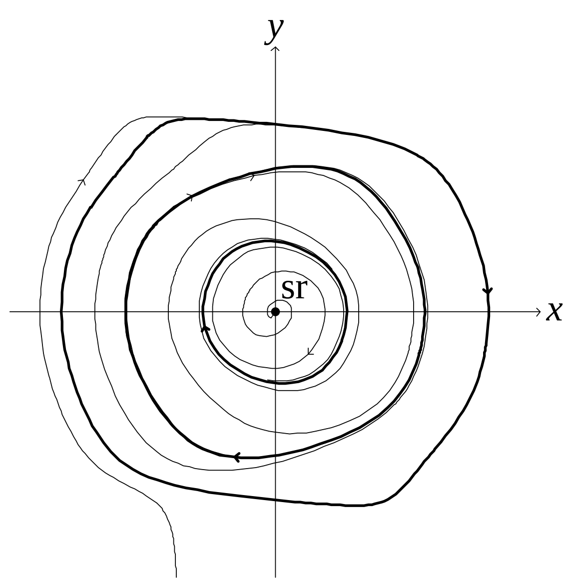

Figure 4: Phase diagram of the aeroelastic galloping model.

In many cases, experts derive asymptotic behavior by assuming that the governing equations are linear over the range of interest and applying the well-developed theory of linear systems. When nonlinearity cannot be ignored, they resort to the mathematical theory of

concept	definition
phase space	geometric representation of state space
trajectory	solution curve in phase space
fixed point	constant solution; a point trajectory
limit cycle	periodic solution; a simple loop trajectory
invariant set	trapping region in phase space
attractor	invariant set to which nearby trajectories converge
attractor basin	invariant set of all trajectories that converge to the attractor

Table 3: Key concepts of phase space analysis.

dynamical systems, which provides qualitative descriptions of the possible attractors, and to numerical software for finding specific attractors and basins.

The SPQR representation does not support asymptotic analysis. Although SPQR can represent fixed points as states whose derivatives all equal 0, it cannot represent the other steady-states. The most important missing case is a limit cycle, such as the waveform of a driven oscillator or the path of a swing pushed by a diligent parent. Transition graphs cannot express the difference between a limit cycle (which yields a happy child), a solution that approaches a fixed point via damped oscillations (an unhappy child), and a solution that moves away from a fixed point via growing oscillations (an endangered child); all three appear as cycles in the graph. For example, all three behaviors in Fig. 5 are consistent with the transition graph of the block and spring (Fig. 3), as discussed in Section 2.5. SPQR lacks any means for representing quasi-periodic and chaotic steady-states, stability, or attractor basins.

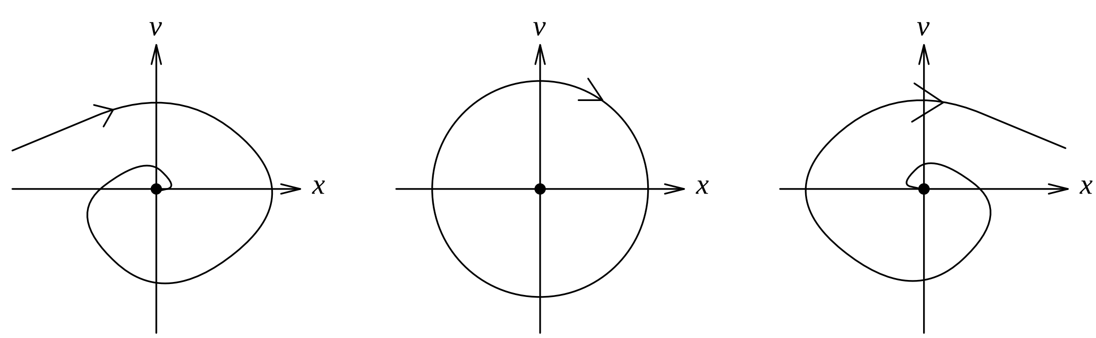

Figure 5: State space depictions of three behaviors consistent with the transition graph of the block and spring: an attracting fixed point (damped oscillation) a limit cycle (periodic oscillation), and a repelling fixed point (growing oscillation).

The SPQR algorithm exacerbates the limitations of the behavioral representation by downplaying linear systems theory, dynamical systems theory, and numerical analysis and by tracing the transient behavior of a system, perhaps forever, rather than directly exploring its asymptotic behavior. It can derive the asymptotic behavior only when every path through the transition graph leads to a fixed point (cf. (Iwasaki and Simon, 1986b)). It cannot handle systems whose graphs contain cycles, such as the block and spring. Some of the SPQR extensions can distinguish limit cycles from spiral fixed points, but only in a few special cases.

We can view SPQR as a limited form of phase space analysis, following Sacks (1990b). The SPQR states of a system translate to rectangular regions (including degenerate rectangles, such as points and lines) in phase space (Fig. 6). The transition graph specifies the realizable transitions between regions. Sacks recasts the SPQR transition test as two algebraic conditions on the equations at the region boundaries. The conditions produce no ambiguous transitions, apply to regions of all shapes, and are testable by an inequality prover.

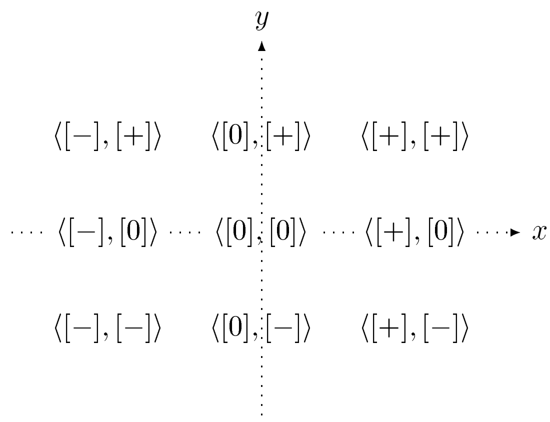

Figure 6: Phase space regions for the SPQR states of the block and spring.

Many of the limitations of SPQR in comparison with expert analysis stem directly from the use of transition graphs over rectangular regions instead of phase diagrams. Rectangular regions cannot represent curved attractors and basins, which are the norm. SPQR could approximate the curved shapes with large numbers of rectangles, but at a high cost in conceptual clarity and in computation. For example, capturing the three limit cycles of the aeroelastic galloping model (Fig. 4) requires hundreds of rectangles. SPQR makes no attempt to approximate attractors, however, and simply partitions phase space according to the signs of the state variables. The resulting transition graphs tend to reflect irrelevant distinctions in the equations, while omitting key distinctions in the solutions. For example, the transition graph of the block and spring (Fig. 3) contains an infinity of paths even though the system has a single asymptotic behavior, since x and v can change signs any number of times as they approach $\langle[0],[0]\rangle$. Yet the transition graph fails to distinguish the true asymptotic behavior from growing, stable, or erratic oscillations.

3.2 Experts hypothesize and revise specific equations

Experts model physical systems with specific differential equations that capture the features of interest. Qualitative physics offers two arguments that specific equations cannot adequately model physical systems whose exact workings are unknown or extremely complex: specific equations must incorporate unwarranted assumptions about the system, hence may yield incorrect inferences; and specific equations may yield no inferences due to limitations on the reasoning or computational abilities of the analyst. Although both problems can occur, the qualitative physics literature contains only anecdotal evidence that they occur in practice. More to the point, it lacks any examples where experts draw incorrect or incomplete conclusions from a specific equation while SPQR does better on a generalized version. On the contrary, SPQR often draws incorrect conclusions or bogs down on problems that experts solve quickly.

The objection to specific equations contradicts a hard-learned lesson of artificial intelligence research: the best way to solve complex, ambiguous problems is often to search through more concrete problems by making and refining reasonable assumptions. This is just what experts do when faced with a system for which they lack precise equations. They iteratively formulate and analyze equations that embody reasonable assumptions (even when they know these assumptions to be false, strictly speaking), compare the solutions with observations, and revise the equations until the discrepancies become insignificant for the problem at hand. Experts start with linear equations because these describe many real-world problems adequately (even if the behavior is not linear at some larger or smaller level of detail), and because one can analyze linear systems quickly and completely. They progress to nonlinear equations when linear equations prove inadequate. Simple equations, such as low-order rational functions, cover essentially everything.

Moreover, unwillingness to risk incorrect conclusions also prevents SPQR from drawing conclusions about reasonable, as opposed to possible, behavior. To ensure the soundness of its conclusions, SPQR must take into account every wild behavior compatible with its general equations. These wild behaviors differ in every way from the normal behaviors, and so preclude sound inference of any normal properties of behaviors.

All this is not to say that experts formulate completely specific equations. They generally formulate parameterized differential equations whose parameters represent approximately known physical constants. They then partition the parameter space into open regions of equivalent systems bounded by *bifurcation curves* where the behavior changes qualitatively (Guckenheimer and Holmes, 1986). Sometimes, experts use other abstractions, for example modeling frictionless mechanisms by *Hamiltonian equations*

$$\dot{x} = \frac{\partial H}{\partial y}$$
$$\dot{y} = -\frac{\partial H}{\partial x}$$

without specifying the Hamiltonian function $H(x, y)$. Our point is that experts immediately reduce any system description to equations that they can hope to analyze, and then revise these equations if they prove unsuitable.

Experts avoid unspecified monotonic functions because these provide few useful inferences. In Appendix B, we prove that every univariate SPQR expression is locally equivalent to one of five constraints on functions of bounded variation: monotone increasing (M^+), monotone decreasing (M^-), positive (P), negative (N), or unconstrained (U). This constrains the expressive power of SPQR equations so that the only first-order SPQR equations are $\dot{x} = M^+(x)$ (positive feedback), $\dot{x} = M^-(x)$ (negative feedback), $\dot{x} = P(x)$ (monotone growth), $\dot{x} = N(x)$ (monotone decay), and $\dot{x} = U(x)$ (no information). This classification result demonstrates that the reduction of (1) to $\dot{v}_y = M^-(v_y)$ presented in Section 2.5 does not represent an isolated simplification. Instead, all first-order problems reduce to one of these five cases. The final three cases are far too general for most reasoning. Thus, first-order SPQR equations can express only two useful models: positive and negative feedback.

Our proof does not apply to multivariate SPQR expressions, such as $M^+(x+y)$, or to SPQR expressions that further constrain their instances, such as M_0^+. But our analysis does extend to higher-order equations where all SPQR expressions are univariate. For example, the equation $\ddot{x} = \Phi(\dot{x}) + \Psi(x)$ has only four instances without P, N, or U functions: $\ddot{x} = M^{\pm}(\dot{x}) + M^{\pm}(x)$. One of these generalizes the spring equation, which still gives SPQR trouble. The others do not appear in the literature.

The weakness of SPQR equations for analyzing dynamical systems does not mean that monotonic functions are useless. Monotonic relationships play an important role in analyzing some static situations, especially in deriving the stability of fixed points (Sacks, 1990b) and in decision analysis, where Wellman (1990) presents a useful, implemented theory of qualitative probabilistic networks based on monotonicity relationships.

3.3 Experts derive qualitative information by numerical analysis

Experts make extensive, informed use of numerical analysis. They find fixed points with algebraic equation solvers, such as Newton-Rhapson iteration, and compute the eigenvalues and eigenvectors of the Jacobian with linear algebra packages. They construct trajectories with differential equation solvers, such as the Runge-Kutta algorithm. More advanced algorithms find saddle manifolds, limit cycles, attractor basin boundaries, and bifurcations.

Qualitative physics eschews numerical analysis, arguing that it provides only reams of numbers, not qualitative information, and that it is prohibitively expensive and unreliable for analyzing realistic systems.[4] This argument ignores the role of expert knowledge in numerical analysis. Experts know that "the purpose of computing is insight, not numbers," as Hamming (1962, Ch. $N+1$) puts it in his classic text on numerical analysis, and they guide their computations accordingly.

Experts infer qualitative information from numerical data based on theoretical and empirical knowledge about plausible outcomes. For example, the phase diagrams above contain much qualitative information even though their details come from numerical analysis. The observed numbers provide strong evidence for the inferences. Although the experts would prefer the infallible support of a proof, they gladly make do with the empirical support of a careful simulation. They infer the existence of a fixed point from an approximate zero of the equations, infer the existence of a limit cycle from a numerically generated trajectory that intersects itself, and derive the other qualitative information analogously.

Experts use their mathematical and domain knowledge about plausible outputs to control the expense and reliability of numerical analysis. Rather than exploring the entire

[4]Forbus (1990, p. 11), for example, argues that "such simulations require immense computational resources. Worse yet, it assumes the existence of a complete set of accurate values for all input parameters. Typically, we just don't have such accurate information, thus forcing us to search a space of parameters corresponding to the ranges the various input parameters may take. This increases the amount of computation even more, making numerical simulation infeasible.

"Even if numerical simulation were technologically feasible, by say shirt-pocket supercomputers, or by allowing rough approximations, it still would be insufficient for our robot. First, we still need to interpret the output of the simulation. A list of numerical values is not the most perspicuous representation of an event. Second, any run of a numerical simulator provides a specific set of predictions ... Often we want to characterize the possibilities that might occur, with some guarantee of completeness."

phase space, they focus on the key trajectories, such as attractors and basin boundaries, and partition the remaining trajectories into equivalence classes. They pick reasonable tolerances based on domain knowledge and on previous outputs. In the hands of experts, numerical analysis of typical ordinary differential equations takes only a few minutes on a scientific workstation. Although some tasks require greater effort, the main problem today is interpretation of the output, not computation speed (cf. (Truesdell, 1984)).

4 Mathematics and expert reasoning

The preceding discussion shows that SPQR lacks the knowledge and methods that experts consider necessary for reasoning about physical systems, with the result that experts routinely understand many systems beyond the ken of SPQR. Though intended to be a theory of reasoning about dynamical systems, SPQR has little knowledge of mathematics (or of physics, for that matter). While SPQR posits that most expert reasoning involves only a few elements of calculus and interval arithmetic, experts are trained to draw on sophisticated mathematical tools that originate in differential topology, dynamical systems theory, ergodic theory, and perturbation theory. The tools apply to abstract equations as well as exact ones and provide qualitative information as well as numerical results. Experts make use of their knowledge either directly, by applying mathematical results to yield the answers of interest, or indirectly, by applying the mathematical results in the design of their algorithms. If one seeks a theory of expert reasoning, it seems reasonable to expect the theory to use at least as much knowledge as the experts do, even if it does not use the knowledge in exactly the same way.

In the following, we examine the nature and role of mathematics in expert reasoning to better understand just why it is so important to expert performance. The simple answer is that mathematics is the best known language for formulating and analyzing models, whether qualitative or quantitative.

4.1 Mathematical concepts serve practical modeling needs

Qualitative physicists believe traditional fields shed little or no light on pre-formal expert reasoning.[5] Qualitative physics rightfully takes automating formulation and revision of models as one of its central problems, and revising a model requires one to make explicit the physical and mathematical assumptions underlying the model so that one may change the faulty assumptions.[6] Standard textbook formulations of physical problems do not make explicit any of these assumptions, and instead presume the reader capable of inferring them. This makes introductory textbook treatments unhelpful for the purposes of qualitative physics.

Advanced mathematics presents an entirely different picture. While physics textbooks concentrate on presenting the "compiled" versions of problems that require the reader to

[5] For example, Weld and de Kleer (1990, p. 2) write "For almost all the examples we considered, the conventional mathematical formulation of physics was useless or unnecessary."

[6] To quote Weld and de Kleer (1990, p. 4) again, "Typically, the [physicist's FORTRAN] program [written to make predictions about some system] has no way to detect when the implicit assumptions under which it was written are violated. Qualitative physics aims to lay bare the underlying intuitions and make them sufficiently explicit, so that they can be directly reasoned with and about."

supply many of the unstated assumptions, the concepts of advanced mathematics provide formal ways for expressing these underlying assumptions. In fact, advanced mathematics concerns itself virtually exclusively with qualitative structures and qualitative properties of behaviors. Mathematics did not arise in a vacuum, but developed its concepts in order to provide the best possible tools for solving practical problems about the design, analysis, and control of physical systems.

Practical problems of design, analysis, and control require one to answer the relevant questions about the actual or intended behaviors of a system using the available knowledge. The most important point about these practical tasks is that they lead, in each domain, to a set of key questions: Will the shell reach its target? Will the tank overflow? Will the sardine and anchovy populations continue to cycle as fishing increases, or will one displace the other entirely? Mathematicians took these non-numerical key questions and worked hard for many years to answer them in the simplest, most general, and most powerful ways possible. They started with everyday physical properties, and then refined, developed, systematized, and isolated these into a number of essentially qualitative properties and notions that constitute the core of mathematics, concepts such as thresholds, boundedness, continuity, stability, and bifurcations.

These key questions, and the concepts essential for answering them, remain the same whether or not one automates the solutions on a computer. Experts addressed these tasks long before the advent of computing machines. The need to solve problems in spite of the severe limitations imposed by manual calculation provided the initial motivation behind the invention of many of the qualitative concepts and rules of modern mathematics. (One should not forget that the word "computer" always referred to a human up until the last few decades.) Even though experts now use computers to help extend their abilities to carry out numerical calculations, they must still rely on their underlying qualitative mathematical knowledge to know when to trust these calculations (Truesdell, 1984).

Let us consider some examples of practical concepts and methods from advanced mathematics.

- One key question in design problems is whether some quantity (pressure in a boiler, say) remains bounded or can increase without limit, so that one can design the device to remain functional and safe. While this question has not been answered for all possible systems, mathematicians have identified a number of qualitative properties of systems which imply boundedness of state values. One of the basic results of topology states that continuous functions over compact regions (such as closed intervals or finite unions of closed intervals) are bounded. Another result states that such functions are also absolutely continuous, that is, can be approximated uniformly over the whole region of interest. Every concept involved in these results is purely qualitative in nature. Indeed, these results apply to functions of any type whatsoever, not just to numerical functions.

- Every electrical engineer makes almost daily use of some of the fundamental results from the theory of functions of a complex variable. One result states that the integral of an analytic function around a closed curve of any shape or size is zero (or more generally, the sum of the multiplicity of the poles of the function enclosed by the curve). Another result states that the integral between any two points is independent of the path taken. Knowledge of these results frees the expert to choose paths of integration so as to make the calculation as easy as possible, or to reduce involved integrals to simple formulas. These results are related to conservation laws, which qualitative physics recognizes to be important, but are more widely applicable and general.

- Engineers model the vast majority of physical systems with linear equations because of their tractability and intuitive properties. The most useful property, called superposition, is that linear combinations of solutions are themselves solutions. Time-invariant systems have particularly simple asymptotic behaviors: all trajectories converge to a global fixed point or all trajectories diverge to infinity.

- Experts ignore implausible cases by assuming *genericity*. A generic property of a set is one that holds for "typical" members. The precise definition varies with the problem domain. The strongest definition is an open dense set of full measure, such as the irrational numbers. We must often settle for weaker definitions, such as "an open dense set" or "a countable intersections of open dense sets in a complete metric space." Experts assume that generic properties hold in their models, unless something special (such as symmetry or conservation laws) indicates otherwise. For example, a *nonhyperbolic* fixed point whose Jacobian has imaginary eigenvalues complicates analysis. Experts can assume that the fixed points of an individual equation are hyperbolic because this property is generic in the set of all differential equations. They cannot assume this when analyzing a one-parameter family of differential equations, but can still assume that the equation has at most one nonhyperbolic fixed point with one imaginary eigenvalue.

We conclude that mathematics provides a rich, well-developed store of *qualitative* concepts and results of proven utility for reasoning about physical systems in all their aspects, not just their dynamics.[7]

4.2 Mathematical reasoning can be automated

Recent developments in qualitative physics support the thesis that mathematics provides the best concepts for model formulation and analysis, in that essentially all the extensions to the basic SPQR algorithm rest upon established mathematics. Kuipers (1988) and Struss (1988) base their criteria for determining that trajectories do not intersect on the Jordan curve theorem. Kuipers (1987) and de Kleer and Bobrow (1984) detect spurious SPQR transitions by reasoning about higher-order derivatives. Sacks (1990b) and Williams (1988) reduce the ambiguity of SPQR arithmetic with symbolic algebra techniques and with interval analysis (Moore, 1979). Iwasaki and Simon (1986a) and Sacks (1990b) base their fixed point analysis on sign stability criteria. Raiman (1986), Weld (1990), and Davis (1987) employ infinitesimals and nonstandard analysis. Doyle and Sacks (1991) base their likelihood predictions for SPQR behaviors on an application of the theory of Markov chains similar to

[7] For other presentations of this idea, see (Browder and Mac Lane, 1978; Jaffe, 1984; Wigner, 1960) and the collections (COSRIMS, 1969; Steen, 1978). For surveys of mathematics, see (Courant and Robbins, 1941) (elementary), (Gårding, 1977; Mac Lane, 1986) (advanced), or (Dieudonné, 1982) (stratospheric). For the qualitative theory of dynamical behaviors, see (Guckenheimer and Holmes, 1986).

that of Hsu (1987). Kuipers's hierarchical models closely resemble the mathematical methods of averaging (Guckenheimer and Holmes, 1986) and of multiple time scales (Brackbill and Cohen, 1985).

But SPQR is a theory of expert reasoning that makes use of no knowledge or methods that are not perfectly intelligible to the educated layman, and its extensions draw on only a minor portion of the knowledge visibly used by experts in their reasoning. If the missing expert knowledge is in fact superfluous, SPQR represents a monumental advance in mathematics and in human inquiry. But the previous discussion shows that SPQR does not obviate other mathematical and scientific reasoning. Thus, the solution to SPQR's limitations involves automating the full range of expert knowledge and methods.

Some basic mathematical knowledge and methods have already been automated with considerable success. For example, Sacks (1990a; 1991) presents an analysis program for one-parameter planar equations that performs at the level of experts by exploiting the mathematical knowledge available to experts. The input is the equations, bounding intervals for the state variables and the parameter, and error tolerances. The program partitions the parameter interval into open subintervals of equivalent behavior bounded by bifurcation points, classifies the bifurcation points, and constructs representative phase diagrams for the subintervals. It constructs the phase diagrams by identifying fixed points, saddle manifolds, and limit cycles and partitioning the remaining trajectories into open regions of uniform asymptotic behavior. It produced the phase diagrams in Figure 4 in a few seconds on a standard scientific workstation. It can solve textbook examples along with problems of practical interest to scientists and engineers, including ones that warranted entire journal articles within the last decade. In other work, Yip (1989) constructs phase diagrams for one-parameter area-preserving planar maps, which model conservative phenomena, such as frictionless bouncing balls. His program treats the equations as a black box for generating trajectories. It classifies trajectories according to their geometry and picks initial points for simulation according to the geometry of existing trajectories. The program performs comparably to experts and has solved an open problem in fluid dynamics. As a third example, Abelson *et al.* (1989) survey other research in automating qualitative analysis via numerical experimentation.

Much work remains to be done. The central problem for qualitative physics must be automating the formulation of models, which is a problem that neither mathematics nor most qualitative physics has addressed.[8] Qualitative physics must also shoulder the burden of adapting mathematical concepts to take computational concerns into account. The precise forms of mathematical concepts useful in engineering reasoning may be different from those useful in mathematics, since the utility of these concepts is different for the two applications. Practical reasoners seek to optimize computational utility, while mathematicians seek to optimize mathematical utility, which takes into account simplicity, beauty, and the power concepts give to human mathematicians in constructing proofs. Optimizing mathematical utility does not necessarily mean optimizing computational utility as well, but modern physics and economics have observed a high correlation. This is not too surpris-

[8] Weld and de Kleer (1990, p. 481) view this as a lacuna of qualitative physics: "By focusing on tasks of analysis and design in the framework of a single, human-provided model, the bulk of work in qualitative physics has finessed what probably is the most important and hardest problem: Constructing an appropriate model."

ing, since most of the qualitative concepts of mathematics were developed to save humans effort in calculation. Finally, mathematics is not yet complete, and problems in qualitative physics may require development of new mathematical concepts.[9]

5 Conclusion

We have argued that the methods that have come to be accepted as the basis for qualitative physics suffer severe limitations in comparison with human experts because they eschew the viewpoint, the tools, and the knowledge of experts, especially the wealth of qualitative concepts and results that modern mathematics provides specifically to facilitate practical and efficient reasoning about the qualitative properties of physical systems. We base our arguments on the following observations:

- The accepted approach, which we have called SPQR, has successfully analyzed only the simplest systems, while routine expert methods succeed on far more complex systems.
- Virtually all of the systems analyzed in the literature reduce to just three equations.
- SPQR equations are far too general for practical use. Experts instead hypothesize and revise specific equations until they obtain equations of adequate accuracy.
- Experts focus on asymptotic behavior, while SPQR focuses on transient behavior.
- Experts derive the behavior of dynamical systems with deep mathematics and extensive numerical analysis, whereas SPQR uses little of either.

To reproduce expert skills, qualitative physics should cast off the fetters of the currently accepted methods and instead seek to exploit modern mathematics (and physics, chemistry, etc.) to the full. It should first seek to automate standard expert knowledge and reasoning methods before deciding to develop entirely different methods. Evidence suggests that the mathematical concepts and results already available suffice to automate substantial amounts of expert reasoning, so neglecting mathematics of proven utility simply slows progress and risks reinventing preliminary versions of established mathematics.

We believe that exploiting standard mathematical concepts and knowledge will also aid qualitative physics in cooperating with other scientific and engineering fields. The current practice, with its implication that mathematics has never addressed the problems of reasoning in any significant way, and with the concomitant "not invented here" requirement that one must abandon established concepts and methods to automate expert reasoning, cannot appear attractive to informed outsiders, and should be abandoned in favor of a more productive spirit of cooperation and building on past discoveries. We should speak of "expert reasoning" (or scientific and engineering reasoning) rather than "qualitative reasoning" when defining the subject, and view qualitative physics as an extension of existing scientific disciplines rather than as an entirely new field of endeavor.

[9] Doyle (1983) uses the term "rational psychology" (in analogy with rational mechanics) for the branch of mathematics aimed at finding the most appropriate concepts for theories in psychology and artificial intelligence.

Acknowledgements

We thank Harold Abelson, Jonathan Amsterdam, Mark Greenstreet, Leo Joskowicz, Benjamin Kuipers, Joseph Schatz, Howard Shrobe, Gerald Sussman, Peter Szolovits, Dan Weld, and Michael Wellman for helpful discussions and comments. Needless to say, however, these people do not necessarily share any of the views we express in this paper, and we accept responsibility for any flaws in this paper. We also thank Guy Steele Jr. for help with Latin. Jon Doyle is supported by the National Library of Medicine through National Institutes of Health Grant No. R01 LM04493. Elisha Sacks is supported by the National Science Foundation under grant No. IRI–9008527 and by an IBM grant.

A Classification of SPQR examples

Table 4 below lists the SPQR analyses contained in Weld and de Kleer (1990). Type 1 is $\dot{x} = M^-(x)$, type 2 is $\ddot{x} = M^-(\dot{x}) + M^-(x)$, and type 3 is a coupled pair of type 1 equations. As noted in the text, SPQR can analyze types 1, 3, and algebraic equations, but cannot fully analyze type 2 or the other types. Mathematicians fully analyzed all these equations before the advent of artificial intelligence (Brauer and Nohel, 1969). Hence, only five SPQR equations support useful inferences about dynamical systems and none support new inferences.

page	**author**	**system**	**type**
99	de Kleer	pressure regulator	1
103	de Kleer	pressure regulator	2
128	Bobrow & de Kleer	pressure regulator	2
139	Williams	RC circuit	1; linear
151	Williams	RR current divider	algebraic
154	Williams	RC current divider	1; linear
157	Williams	RC high-pass filter	1; linear
160	Williams	RCRC ladder	3; linear
161	Williams	Wheatstone bridge	algebraic
167	Williams	n-Mosfet	other
196	Forbus	U-tube	1
198	Forbus	heat flow	1; linear
207	Forbus	spring	2
211	Forbus	W-tube	3
231	Forbus	heat flow	1; linear
237	Kuipers	U-tube	1
244	Kuipers	flying ball	algebraic
249	Kuipers	spring	2
262	Kuipers & Chiu	cascaded tanks	1
262	Kuipers & Chiu	coupled tanks	3
270	Lee & Kuipers	spring	2
273	Struss	spring	2

308	Dormay & Raiman	pressure regulator	2
313	Williams	U-tube	1
318	Raiman	colliding masses	algebraic
359	Williams	spring	2
398	Weld	spring	2
399	Weld	heat exchanger	1; linear
409	Weld	Wheatstone bridge	algebraic
420	Weld	heat exchanger	1; linear
427	Davis	block	2
438	Dague & Raiman	voltage follower	other
444	Forbus	heat flow	1; linear
531	Kuipers	water balance	1
531	Kuipers	sodium balance	1
632	Iwasaki & Simon	evaporator	1; linear
642	Iwasaki & Simon	pressure regulator	2

Table 4: Classification of systems analyzed in Weld and de Kleer (1990).

B Classification of SPQR equations

The small repertoire of useful SPQR equations reflects their limited expressive power. We use a standard result about monotonic functions to make this claim precise. Given a function f defined on the interval $[a, b]$ and a subdivision $a = x_0 < x_1 \cdots < x_n = b$ of $[a, b]$, define

$$t = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|.$$

The function f is of *bounded variation* over $[a, b]$ if the supremum of t over all subdivisions of $[a, b]$ is finite.

Lemma 1 *Let Φ be of bounded variation on $[a, b]$. There exist $f, g \in M^+$ such that $\Phi = f - g$ on $[a, b]$.*

Proof: We can find monotonic functions $\hat{f}$ and $\hat{g}$ such that $\Phi = f - g$ on $[a, b]$ by a standard theorem (Royden, 1968, p. 100). Setting $f = \hat{f} + x$ and $g = \hat{g} + x$ guarantees $f, g \in M^+$. $\square$

We use this result to characterize the expressive power of univariate SPQR expressions. Define an SPQR expression in x as $M^+(x)$, $M^-(x)$, or any expression $\alpha + \beta$, $-\alpha$, $\alpha\beta$, $1/\alpha$, $M^+(\alpha)$, or $M^-(\alpha)$ in which α and β are SPQR expressions in x. Every SPQR expression is of bounded variation under the standard SPQR assumptions (Kuipers, 1986). Define the extended SPQR values on $[a, b]$ as $\{M^+, M^-, P, N, U\}$ where M^+ and M^- are strictly increasing and decreasing continuous functions on $[a, b]$ and P, N, and U denote any positive, any negative, or simply any function of bounded variation over $[a, b]$.

Theorem 2 *Let Φ be an SPQR expression in x defined on $[a, b]$. There exists a subdivision $a = x_0 < x_1 \cdots < x_n = b$ of $[a, b]$ such that Φ is equivalent to an extended SPQR value on each subinterval $[x_{i-1}, x_i]$.*

Proof: The proof is by induction on the nesting level of Φ. If $n = 1$, the only possible expressions are $M^+(x)$ and $M^-(x)$, which both satisfy the condition on $[a, b]$. Assuming $n - 1$, the proof for $n > 1$ is by cases. The cases apply to the intervals of the $n - 1$ subdivision. If one argument is U, it absorbs the other argument and U is the result. The values P, N, and U are their own reciprocals. We now consider the other cases.

$M^\pm + P = U$: We can write the expression $M^+ + P$ as $M^+ + (P + k)$ for any constant k by rule 5 of Table 1. We can write any f on $[a, b]$ as $P + k$ with k the absolute value of the infimum of f on $[a, b]$. Hence, $M^+ + (P + k) = M^+ + U$, which equals U as shown above. We reduce $M^- + P$ to $-M^+ + P$ and so to U as well. The cases $M^\pm + N = U$ are analogous.

$M^\pm(P) = U$: We obtain any f as the result of $M^+(P)$ by setting M^+ to $\lambda x.x + k$ with k as above. We reduce $M^-(P)$ to $-M^+(P)$ and so to U as well. The cases $M^\pm(N) = U$ are analogous.

$M^\pm \times M^\pm$: Let $f, g \in M^+$. Divide the interval $[a, b]$ into four subintervals (some possibly empty): (i) $f, g > 0$, (ii) $f, g < 0$, (iii) $f > 0; g < 0$, and (iv) $f < 0; g > 0$. The result is M^+ on (i) by rule 1 of Table 1 and M^- on (ii) by rule 2 because $(fg)' = f'g + fg'$ and $f', g' > 0$. On interval (iii), we have

$$fg = -e^{\log(-fg)} = -e^{\log(f)+\log(-g)} = -e^{M^+ + M^-} = -e^U = -P = N.$$

The third equality holds because any function can be written as the log of a positive function; the fourth holds by Lemma 1; and the fifth holds because any positive function can be written as the exponential of a function. Interval (iv) yields P analogously. The remaining cases of $M^\pm \times M^\pm$ are similar.

$M^+ \times P$: Let $f \in M^+, g \in P$, and divide the interval $[a, b]$ into two subintervals corresponding to $f < 0$ and $f > 0$ (one of these may be empty). The product is N in the subinterval where $f < 0$, as it generalizes interval (iii) of the previous case, and is P in the $f > 0$ interval, as g absorbs f. The cases $M^- \times P$ and $M^\pm \times N$ reduce to this case or its negation.

The remaining cases follow from direct applications of Lemma 1 and the rewrite rules in Table 1. □

As mentioned in the text, this result means that the only first-order SPQR equations are $\dot{x} = M^+(x)$ (positive feedback), $\dot{x} = M^-(x)$ (negative feedback), $\dot{x} = P(x)$ (monotone growth), $\dot{x} = N(x)$ (monotone decay), and $\dot{x} = U(x)$ (no information). The final three cases are far too general for most reasoning. Thus, first-order SPQR equations can express only two useful models: positive and negative feedback. This conclusion extends to higher-order equations where all SPQR expressions are univariate. For example, the equation $\ddot{x} = \Phi(\dot{x}) + \Psi(x)$ has only four instances without P, N, or U functions: $\ddot{x} = M^\pm(\dot{x}) + M^\pm(x)$. Only one of these appears in the literature, and it gives SPQR trouble.

References

Abelson, Harold; Eisenberg, Michael; Halfant, Matthew; Katzenelson, Jacob; Sacks, Elisha; Sussman, Gerald Jay; Wisdom, Jack; and Yip, Kenneth 1989. Intelligence in scientific computing. *Communications of the ACM* 32(5):546–562.

Abraham, R. and Marsden, J. E. 1978. *Foundations of Mechanics.* Benjamin, Reading, MA, second edition.

Arnold, V. I. 1984. *Mathematical Methods of Classical Mechanics.* Springer-Verlag, New York.

Benson, Sidney W. 1982. *The Foundations of Chemical Kinetics.* Krieger, Malabar, FL, second edition.

Brackbill, J.U. and Cohen, B.I. 1985. *Multiple Time Scales.* Academic Press, New York.

Brauer, Fred and Nohel, John A. 1969. *The Qualitative Theory of Ordinary Differential Equations.* W.A. Benjamin, Inc., New York.

Browder, Felix E. and Mac Lane, Saunders 1978. The relevance of mathematics. In Steen, Lynn Arthur, editor 1978, *Mathematics Today: Twelve Informal Essays.* Springer-Verlag, New York. 323–350.

Chua, Leon O.; Desoer, Charles A.; and Kuh, Ernest S. 1987. *Linear and Nonlinear Circuits.* McGraw-Hill, New York.

Clark, Colin W. 1976. *Mathematical Bioeconomics: The Optimal Management of Renewable Resources.* Wiley, New York.

COSRIMS, 1969. *The Mathematical Sciences: A Collection of Essays.* MIT Press, Cambridge, MA. Edited by the Committee on Support of Research in the Mathematical Sciences (COSRIMS) with the collaboration of George A. W. Boehm.

Courant, Richard and Robbins, Herbert 1941. *What is Mathematics?* Oxford University Press, London.

Crawford, James; Farquhar, Adam; and Kuipers, Benjamin 1990. QPC: a compiler from physical models into qualitative differential equations. In *Proceedings of the National Conference on Artificial Intelligence.* 365–372.

Dague, P.; Devès, P.; and Raiman, O. 1987. Troubleshooting: When modeling is the trouble. In *Qualitative Physics Workshop Abstracts.*

Davis, Ernest 1987. Order of magnitude reasoning in qualitative differential equations. Technical Report 312, New York University Computer Science Department.

de Kleer, J. 1977. Multiple representations of knowledge in a mechanics problem-solver. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence.* Reprinted in (Weld and de Kleer, 1990).

de Kleer, Johan and Bobrow, Daniel G. 1984. Qualitative reasoning with higher-order derivatives. In *Proceedings of the National Conference on Artificial Intelligence.* American Association for Artificial Intelligence. 86–91.

de Kleer, Johan and Brown, John Seely 1984. A qualitative physics based on confluences. *Artificial Intelligence* 24:7–83.

de Kleer, Johan and Brown, John Seely 1986. Theories of causal ordering. *Artificial Intelligence* 29:33–61.

Dieudonné, Jean 1982. *A Panorama of Pure Mathematics, As Seen by N. Bourbaki.* Academic Press, New York. Translated by I. G. Macdonald.

Doyle, Jon 1983. What is rational psychology? toward a modern mental philosophy. *AI Magazine* 4(3):50–53.

Doyle, Jon and Sacks, Elisha P. 1991. Markov analysis of qualitative dynamics. *Computational Intelligence* 7(1). To appear.

Etkin, Bernard 1972. *Dynamics of Atmospheric Flight.* Wiley, New York.

Falkenhainer, Brian and Forbus, Kenneth 1988. Setting up large-scale qualitative models. In *Proceedings of the Seventh National Conference on Artificial Intelligence.* 301–306.

Forbus, Kenneth D. 1984. Qualitative process theory. *Artificial Intelligence* 24:85–168.

Forbus, Kenneth D. 1990. Qualitative physics: Past present and future. In Weld and de Kleer (1990). 11–39.

Gårding, Lars 1977. *Encounter With Mathematics.* Springer-Verlag, New York.

Guckenheimer, John and Holmes, Philip 1986. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, New York.

Hamming, R. W. 1962. *Numerical Methods for Scientists and Engineers.* McGraw-Hill.

Hirsch, Morris W. and Smale, Stephen 1974. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press College Division, Orlando, Florida.

Hsu, Chieh Su 1987. *Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems*, volume 64 of *Applied Mathematical Sciences.* Springer-Verlag, New York.

Iwasaki, Yumi and Simon, Herbert A. 1986a. Causality in device behavior. *Artificial Intelligence* 29:3–32.

Iwasaki, Yumi and Simon, Herbert A. 1986b. Theories of causal ordering: Reply to de Kleer and Brown. *Artificial Intelligence* 29:63–72. Reprinted in (Weld and de Kleer, 1990).

Jaffe, Arthur M. 1984. Ordering the universe. *SIAM Review* 26(4):473–500.

Kuipers, Benjamin 1984. Commonsense reasoning about causality: Deriving behavior from structure. *Artificial Intelligence* 24:169–204.

Kuipers, Benjamin 1985. Qualitative simulation in medical physiology: A progress report. TM 280, Massachusetts Institute of Technology, Laboratory for Computer Science, 545 Technology Square, Cambridge, MA, 02139.

Kuipers, Benjamin 1986. Qualitative simulation. *Artificial Intelligence* 29:289–338.

Kuipers, Benjamin 1987a. Abstraction by time-scale in qualitative simulation. In *Proceedings of the National Conference on Artificial Intelligence*. American Association for Artificial Intelligence. 621–625.

Kuipers, Benjamin 1987b. New reasoning method for artificial intelligence in medicine. *International Journal of Man-Machine Studies* 26:707–718.

Kuipers, Benjamin and Berleant, Daniel 1988. Using incomplete quantitative knowledge in qualitative reasoning. In *Proceedings of the National Conference on Artificial Intelligence*. American Association for Artificial Intelligence.

Kuipers, Benjamin and Chiu, Charles 1987. Taming intractible branching in qualitative simulation. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*. International Joint Conference on Artificial Intelligence. 1079–1085. Reprinted in (Weld and de Kleer, 1990).

Lee, Wood and Kuipers, Benjamin 1988. Non-intersection of trajectories in qualitative phase space: A global constraint for qualitative simulation. In *Proceedings of the Seventh National Conference on Artificial Intelligence*. American Association for Artificial Intelligence.

Mac Lane, Saunders 1986. *Mathematics: Form and Function*. Springer-Verlag, New York.

Marsden, J. E. and Hughes, T. J. R. 1983. *Mathematical Foundations of Elasticity*. Prentice-Hall, Engelwood Cliffs, NJ.

Mavrovouniotis, Michael and Stephanopoulos, George 1987. Reasoning with orders of magnitude and approximate relations. In *Proceedings of the National Conference on Artificial Intelligence*. American Association for Artificial Intelligence. 626–630.

Molle, David Dalle 1989. Qualitative simulation of dynamic chemical processes. Technical Report AI89-107, Artificial Intelligence Laboratory of The University of Texas at Austin.

Moore, Ramon E. 1979. *Methods and Applications of Interval Analysis*. SIAM Studies in Applied Mathematics. SIAM, Philadelphia.

Parker, T. S. and Chua, L. O. 1989. *Practical Numerical Algorithms for Chaotic Systems*. Springer-Verlag, New York.

Prigogine, Ilya 1980. *From Being to Becoming: Time and Complexity in the Physical Sciences*. W. H. Freeman, San Francisco.

Raiman, Olivier 1986. Order of magnitude reasoning. In *Proceedings of the National Conference on Artificial Intelligence*. American Association for Artificial Intelligence. 100–104.

Royden, H. L. 1968. *Real Analysis.* Macmillan Publishing Inc., New York, second edition.

Sacks, Elisha P. 1990a. Automatic phase space analysis of dynamical systems. *Computing Systems in Engineering* 1(2-4):607–614.

Sacks, Elisha P. 1990b. A dynamic systems perspective on qualitative simulation. *Artificial Intelligence* 42(2-3):349–362.

Sacks, Elisha P. 1991. Automatic analysis of one-parameter planar ordinary differential equations by intelligent numerical simulation. *Artificial Intelligence* 48(1):27–56.

Steen, Lynn Arthur, editor 1978. *Mathematics Today: Twelve Informal Essays.* Springer-Verlag, New York.

Struss, Peter 1988. Global filters for qualitative behavior. In *Proceedings of the Seventh National Conference on Artificial Intelligence.* American Association for Artificial Intelligence.

Thompson, J. M. T. and Stewart, H. B. 1986. *Nonlinear Dynamics and Chaos.* John Wiley and Sons, New York.

Truesdell, Clifford 1984. The computer: ruin of science and threat to mankind. In *An Idiot's Fugitive Essays on Science: Methods, Criticism, Training, Circumstances.* Springer-Verlag, New York. 594–631.

Weld, Daniel S. 1990. *Theories of Comparative Analysis.* MIT Press.

Weld, Daniel S. and de Kleer, Johan, editors 1990. *Readings in Qualitative Reasoning about Physical Systems.* Morgan Kaufman, San Mateo, Ca.

Wellman, Michael P. 1990. *Formulation of Tradeoffs in Planning Under Uncertainty.* Pitman and Morgan Kaufmann.

White, Frank M. 1986. *Fluid Mechanics.* McGraw-Hill, New York, second edition.

Wigner, Eugene P. 1960. The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13:1–14.

Williams, Brian C. 1984. Qualitative analysis of MOS circuits. *Artificial Intelligence* 24:281–346. Reprinted in (Weld and de Kleer, 1990).

Williams, Brian C. 1988. Minima: A symbolic approach to qualitative algebraic reasoning. In *Proceedings of the National Conference on Artificial Intelligence.* American Association for Artificial Intelligence. 264–269.

Yip, Kenneth Man-kam 1989. Kam: Automatic planning and interpretation of numerical experiments using geometrical methods. AI-TR 1163, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA, 02139.

Reset reproduction of article published in *Computational Intelligence*, Vol. 8, No. 2 (May 1992), pp. 326–335. Reprinted July 1994.

Epilegomenon

Elisha P. Sacks
Department of Computer Science, Princeton University, Princeton, New Jersey 08544

Jon Doyle
Laboratory for Computer Science, Massachusetts Institute of Technology
545 Technology Square, Cambridge, Massachusetts 02139

We are deeply indebted to the commentators for the hard work they have applied to responding to our paper. Neither their efforts nor our own in responding to their comments have been as enjoyable as one might wish. Our paper apparently persuaded some commentators that we hold positions that we thought we had explicitly denied, positions we denied for many of the same reasons as have the commentators. Though we expected disagreements, we were very surprised by some of the points actually taken to represent major controversies. We did not anticipate the variety of ways in which placing different emphases on our words yields unintended interpretations. We would have chosen different words in some cases, and probably a different organization for the text, had we foreseen these unintended interpretations. We regret that our paper is not the one we wish we had written, and regret that our tongue-in-cheek title may have given offense, for none was intended. To make amends, we will attempt in this afterword to express our intentions more clearly. Rather than attempt a comprehensive response to every point raised by the commentators, we restate our major points and principal arguments, saying exactly what we intended to say in our paper, but in a form we hope will be less conducive to misunderstanding.

1 Our intention

Our intention in writing the paper was to encourage work on the problems of qualitative physics by calling attention to the overwhelmingly *qualitative* nature of the many concepts and results modern mathematics provides for representing knowledge about the physical world, and by proposing that vigorous exploitation of these concepts and results in automated reasoners promises the most *direct* path to mechanizing scientific and engineering reasoning.

In explicitly focusing on the qualitative physics project of automating scientific and engineering reasoning about physical systems, we did not intend to equate qualitative physics with this one project or to denigrate its other projects, such as elucidating and mechanizing commonsense and causal reasoning about the world. We view these other projects as both interesting and important, and do not believe that any success or failure of some method in automating scientific and engineering reasoning necessarily entails success or failure in these other projects. Nor did we intend to equate qualitative physics with any specific reasoning task like prediction, or with the set of concepts, representations, and algorithms we call SPQR. Most importantly, we believe that while qualitative physics has yet to produce an automated scientist or engineer, this simple fact says nothing about the ultimate success or value of the field, since few grand human endeavors ever achieve their aims in short order.

We intended all of our discussion concerning "experts" to refer to scientists and engineers, and introduced the term in reference to Weld and de Kleer's stated aim of constructing artificial scientists and artificial engineers. We intended our generic statements about what experts do to refer to what typical experts typically do, and not to things that all experts, without exception, do in all cases, without exception. In particular, we intended our discussion of expert behavior to concern the responsible practices of the broad spectrum of scientists and engineers, and did not intend to restrict attention to the special practices of academic dynamicists or any other subpopulation.

We of course intended our paper to speak to AI researchers already working in qualitative physics. But as we expected that they were already cognizant of many of the issues we discuss, even if they might not share our opinions on these issues, we viewed as even more important the task of alerting students contemplating work in the field to the existence of many useful concepts not commonly encountered in computer science educations. We also hoped to speak to mathematicians, scientists, and engineers looking for opportunities to apply their expertise in a new area. For each of the audiences, we sought to encourage more work on automating qualitative reasoning about physical systems, not to discourage anyone from the enterprise.

2 Our thesis

The main thesis of our paper is that mathematics provides an *extensive*, *well-developed*, *formal*, and *qualitative* language for describing and characterizing the structure and behavior of a very wide variety of systems and for formulating and solving a wide variety of practical problems arising in prediction, design, diagnosis, and control.

In making this point, we intended to follow a tradition of papers in mathematics and physics (Browder and Mac Lane, 1978; COSRIMS, 1969; Jaffe, 1984; Mac Lane, 1986; Truesdell, 1984; Wigner, 1960) that attempt to counter the attitude (widespread throughout civilization) that mathematics concerns only numbers and equations (or worse yet, only numerical solution of ordinary differential equations). In fact, mathematics is not identical with numbers and equations any more than computer science is identical with bits and microprocessors. The intellectual cores of both fields concern qualitative, non-numerical structures for describing real and imaginary systems, situations, and processes, and these structures underlie most of the successful techniques employed by scientists and engineers. We did not intend to suggest that scientists and engineers always consciously or knowingly exploit these qualitative concepts, but only that these concepts provide the basis for many of the techniques successfully employed by scientists and engineers: techniques they employ successfully even if they lack knowledge of the underlying formal theories (but even more successfully when they do know these theories).

The main moral we drew from this thesis is that the most expeditious way of automating scientific and engineering reasoning involves automating mathematical concepts and knowledge directly, given the proven utility of mathematics in aiding human reasoning about both expert and everyday situations, and given that most mathematical concepts and results are already formalized. Creating artificial scientists and engineers requires more than this, of course; we undoubtedly will need to invent some new mathematics along the way (cf. (Doyle, 1983)). For example, mathematics does not provide a language for describing shapes and regions that meets the demands of some practical applications. Practical reasoning tasks often need to characterize shapes and regions by their use, rather than by their expression (in, for example, landmark values of algebraic and trigonometric expressions). We expect that qualitative physics work on kinematics (e.g., Faltings, Joskowicz) and artificial intelligence work on vision constitute steps in this direction. More immediately, though, the point of greatest leverage for AI is in automating the process of model formulation. Even though mathematics provides some of the best concepts and analytical tools for use in constructing models, it provides no formal procedures for using its concepts and tools in actually carrying out such constructions. In contrast, we view many works in AI as developing automatic procedures for formulating problems and models over a variety of representations. These procedures often embody important ideas entirely independent of any use of SPQR concepts and representations.

While we focused our discussion on the automation of mathematical concepts and knowledge, one cannot hope to construct an automated scientist or engineer without also automating non-mathematical knowledge from physics and other worldly subjects. This truism formed the basis for our suggestion of changing talk about "qualitative reasoning" to talk about "expert reasoning" when discussing automated scientific and engineering reasoning. As qualitative physics has observed, however, textbooks in physics and other fields do not always present their qualitative knowledge in qualitative form. (Texts in rational mechanics, however, present an instructive counterexample. See, for example, (Truesdell, 1977).) We thus focused on automating the qualitative concepts of mathematics as a step enabling a more perspicuous formalization of physics and other subjects, but in no way intended to suggest that automating mathematics alone suffices to automate scientific and engineering reasoning.

In making these points, we did not mean to imply that qualitative physics attaches no importance to mathematics. Qualitative physics continues to incorporate additional mathematical ideas beyond those represented in SPQR. Taken together over time, these increments will, as Archimedes suggested, yield success. We nevertheless urged changing current practice in the field because, in talking with people outside the field, we found it conveys to outsiders the impression of a field suspicious of mathematics and externally developed concepts. (We note that AI generally has often conveyed a similar impression: cf. (Doyle, 1988).) We did not believe that that superficial impression accurately reflects the more complex attitudes of qualitative physics researchers concerning the costs and benefits of mathematics. But accurate or not, the impression needlessly damages reputations and impedes cooperation with outsiders who feel offended by the field as they perceive it.

While we observe reasonable differences of opinion in the commentaries on the comparative utility of deliberate automation of mathematical concepts and knowledge, we see no substantial disagreements on the importance of model formulation. We believe almost everyone in qualitative physics also views model construction as a central problem, if not the central problem, of the field, whether or not they choose to work on it themselves. In echoing Weld and de Kleer's observation that the the field has paid less attention to model construction than it deserves, we in no way intended to imply that the field totally neglects the problem. To the contrary, we believed that this is where the field shines brightest, and the thrust of our main point was to help it shine even brighter by exploiting qualitative mathematical concepts more rapidly. Our intent was not to propose or critique any specific procedures for constructing models, only to urge that the models constructed make use of the relevant mathematical concepts.

3 Our arguments

To support the thesis that direct automation of the qualitative concepts of mathematics provides the most expeditious way of automating scientific and engineering reasoning, we provided arguments both for the efficacy of these concepts and for their advantages over the alternatives.

3.1 The efficacy of mathematics

We presented three arguments for the efficacy of mathematical concepts. Two of these concern its epistemological adequacy (in the sense of McCarthy and Hayes), looking at how mathematical concepts were developed in order to describe the world (the design argument), and looking at some examples of how mathematical theories provide concepts appropriate to making important qualitative distinctions (the inductive argument). The third concerns the heuristic adequacy of mathematics, looking at how some mathematical concepts have entered into automated reasoners (the pragmatic argument).

3.1.1 The design argument

In the design argument, we observed that mathematicians did not generally invent their concepts and results as abstract playthings, but as practical means of reaching targeted conclusions about expert and commonsense situations. Mathematicians approached practical problems by determining both the minimal information necessary to reach particular conclusions and the most useful abstractions for each task. Comparisons of the utility of abstractions, while rarely explicit, involved computational difficulty as well as esthetic criteria like simplicity and beauty since mathematicians knew they had to compute the answers themselves. Traditional computational complexity criteria were, in some ways, even more harsh than those applied today since, as we noted in our paper, "computers" were human beings until very recently (cf. (Truesdell, 1984)).

Those mathematicians, such as G. H. Hardy, who publicly delighted in working on beautiful theories they assumed would be forever totally impractical, would be horrified to learn just how regularly modern science finds parts of mathematics to be exactly what it needs to formulate and solve some intensely practical problem. In Hardy's case, the number theory he so admired for its impracticality now serves as the basis for work on reliable and secure communications. Even category theory, long derided as "abstract nonsense" by mathematicians themselves, now justifiably plays increasingly important roles in the design and semantics of programming languages. These constitute just two of the newest cases of what Wigner (1960) called "the unreasonable effectiveness of mathematics" and what Jaffe (1984) identified as mathematics' concern with "ordering the universe".

The reason mathematics proves so successful in practice, whether by intent or not, is because its main concern is to find the most appropriate representations and rules for reasoning about different subjects for different purposes. Consider the attributes of good reasoning methods and representations stated in the responses by Joskowicz and the Xerox SERA group. Joskowicz suggests evaluating methods with respect to the criteria of ontological adequacy, inferential adequacy, inferential parsimony, coverage, and computational efficiency. Williams, Shirley, Raiman, Falkenhainer, de Kleer, and Bobrow paraphrase Winston to write "Good representations facilitate problem solving. They make important things explicit and expose natural constraints. They are complete, concise, transparent, facilitate computation and suppress detail." These two sets of attributes say essentially the same thing. But more importantly, both describe the essence of the mathematical method in formulating, formalizing, and investigating a subject (cf. (Doyle, 1983)). In other words, mathematics has been working on finding good ways of representing and reasoning about the world for thousands of years, and has, together with natural philosophy, constituted the field of "qualitative reasoning about physical systems" for most of the time prior to the rise of artificial intelligence.

3.1.2 The inductive argument

The inductive argument simply points to a variety of examples of qualitative mathematical theories and the useful concepts and methods they provide, leaving the reader to infer that the theories, concepts, and methods not mentioned explicitly are similarly useful in treating other problems. We thought this argument preferable to a complete survey of mathematics only because few readers would wish to read such a proof by exhaustion (literally, given the scope of mathematics) even if we believed ourselves competent to write it. We chose our examples to reflect a bit of the diversity of mathematical concepts, including dynamics, topology, analytic functions, functional analysis, linear relationships, and measure theory. No short list like this can convey much, and perhaps some other selection of theories would have illustrated our thesis better.

As our main example, we delved deepest into the mathematical theory of dynamics, both because of the prominence of dynamical problems in the literature of qualitative physics and because of our familiarity with the theory. We did not intend to equate the overall problem of reasoning about physical systems with the subtask of predicting the behavior of dynamical systems, nor did we mean to imply that all reasoning about physical systems involves the concepts of dynamics. More importantly, we in no way intended our claims for the utility of mathematical concepts to be restricted to dynamics or the few other subjects explicitly mentioned in our paper, as these are just some examples among many. Instead, we believe mathematics supplies the best known formal concepts for understanding all aspects of the world, whether concerning structure or process, statics or dynamics, and whether applied to tasks of prediction, design, diagnosis, or control.

3.1.3 The pragmatic argument

The pragmatic argument points to the effectiveness of mathematical concepts in qualitative physics when the relevant concepts are applied in appropriate ways. Some problems call for exactly the concepts embodied in SPQR, hence the apt success of SPQR on them. Other problems call for additional concepts, and the extensions to SPQR add in some of these. The concepts embodied in SPQR are inappropriate to other problems, but systems automating the knowledge relevant to these problems (e.g., KAM (Yip, 1989) and POINCARE (Sacks, 1991)) have exhibited initial successes.

3.2 The advantages of mathematics

We provided three arguments for the advantages of direct automation of mathematics as a means to automating scientific and engineering reasoning. The first argument simply repeats the design argument for efficacy above with a different emphasis: given the centuries of concentrated effort by thousands of mathematicians on finding the best formalizations of worldly structures and phenomena, it seems unlikely that significantly better ways of formalizing these same phenomena will be found without comparable effort. We recognize that it occasionally may be easier to develop adequate formalizations from scratch than to find and apply the relevant extant mathematics; but we expect these cases to be the exception rather than the rule, and believe that the intellectual and social benefits of vigorous, deliberate exploitation of mathematics outweigh the costs. The second argument is mainly indirect, pointing out the limitations of the available alternatives: using SPQR alone, and using incremental extensions of SPQR. The third argument points out some specific advantages of direct automation over incremental extensions of SPQR.

3.2.1 Limitations of SPQR

Our claim concerns identifying the best path to future capabilities, not the best existing systems. Thus the question of how to automate scientific and engineering reasoning most rapidly would be moot if some existing automated system provided a reasonable approximation of the broad spectrum of reasoning abilities of human scientists and engineers. We know of no one who believes any extant system does this, but we examined the progress achieved by systems based on SPQR as a way of both verifying this assessment and identifying the strengths and weaknesses of its approach as a path to creating an automated scientist or engineer. We found that extant systems based on SPQR provide an understanding of some mathematically simple systems, but fail on other simple systems easily understood completely by scientists and engineers. Some of these failures are overcome by some of the extensions to SPQR, and some by using completely different sets of abstractions, but the goal of constructing an automated scientist or engineer remains to be achieved. In examining the failures, we found severe limitations having to do with the expressiveness of the language and representations SPQR provides, in addition to computational limitations due to the algorithms it employs. Some of these limitations were found by analyzing its language and representations on their own, and some by comparing the SPQR approach with some approaches employed by human scientists and engineers. In making these comparisons, we in no way intended to set up a competition between the current performance of systems based on SPQR and humans, or between systems based on SPQR and automated systems based on any other approach. We intended only to compare future prospects. Given the short history of the field, we viewed current performance as one of the least informative indicators of future prospects.

Inexpressiveness: We found the conceptual language of SPQR to be highly limited in its ability to draw some distinctions important for reasoning about the behavior of systems. Its language represents some systems with qualitatively different behaviors by the same generalized differential equation, such as systems described by the differential equations $\dot{x} = x$ and $\dot{x} = x^3$. It also represents other systems with the same qualitative behaviors by different generalized equations. We presented a set of simple rewrite rules, successors to similar rules first observed by Kuipers, that may be applied in short order to reduce complex sets of generalized equations to simpler ones describing the same behaviors. We also drew on a standard theorem concerning monotone functions to provide an indication of the inherent inexpressiveness of generalized equations based on monotonicity relationships. Finally, we presented a summary of some of the concepts underlying the mathematical theory of qualitative dynamics to illustrate the difficulty of expressing some important distinctions of proven utility with SPQR concepts and representations.

In finding SPQR's concepts and representations inexpressive, we did not intend to suggest that understanding a system requires making unambiguous predictions. As noted by several commentators, all abstractions hide some distinctions and highlight others; when choosing an abstraction for a particular reasoning task, one asks whether it hides all the unimportant distinctions and highlights all the important distinctions. Every abstraction, therefore, introduces a deliberate ambiguity about irrelevant distinctions. The ambiguities inherent in SPQR concepts form limitations only because SPQR analyses remain ambiguous about important qualitative properties of systems concerning asymptotic behaviors, while the corresponding qualitative concepts from mathematics make more useful qualitative distinctions among asymptotic behaviors.

We intended our discussion of the inability of SPQR concepts to distinguish qualitatively different asymptotic properties as just one example of the limitations of its conceptual language; it lacks the appropriate concepts to characterize other qualitative properties as well, even when asymptotic behavior is not important. We did not intend to suggest that scientists and engineers never find transient behavior important, only that most transient analyses presuppose some reference to an equilibrium state or cycle. The practical problem need not be one of determining the precise nature of the baseline state or attractor; that may be be already known, easily determined, or even difficult or infeasible to determine. But most analyses depend on some knowledge about the reference behavior (as in Wellman's distinction between absolute and relative analyses), thus making a practical concern out of the expressive limitations of SPQR for characterizing the qualitative properties of asymptotic behavior.

This examination of the expressiveness of SPQR casts the past successes of SPQR in a new light, since the reduction rules quickly collapse the representations of the seeming wide variety of systems analyzed by SPQR to just a handful of simple generalized equations. We surveyed the 55 papers in the Weld and de Kleer collection, which we took to present the state of the art through 1989, examining all analyses, including those based on extensions to SPQR. We found that after applying the rewrite rules all but two of the successful analyses involve three simple equations.

Lack of routine methods: While the expressive limitations of SPQR prevent it from exploiting some information relevant to understanding behaviors qualitatively, the algorithms it offers do not support some of the methods routinely used by human scientists and engineers in reaching a qualitative understanding of a system, namely the method of modeling systems using parameterized equations, and the method of determining qualitative properties through selected numerical experiments.

Scientists and engineers, like other humans, and like AI systems generally, do not always restrict their reasoning to sound, deductive elaboration of only those consequences strictly entailed by what they know; instead, they ordinarily make reasonable assumptions and guesses, and revise these as they see fit. Only in special cases involving great stakes do they even attempt to consider all and only the logically possible consequences of their hypotheses, and then usually only after some disaster awakens them to the danger. We see no reason to reduce the utility of qualitative physics on the broad range of mundane, small-stakes applications by restricting it to use only methods that provide guarantees appropriate to the narrower range of exotic, great-stakes applications.

One of the most common assumptions scientists and engineers make is to model systems in terms of generalized equations. But the generalized equations they use take a very different form than those employed in SPQR. Rather than use equations over monotone functions, they use parametric and piecewise-linear equations of varying degrees of complexity. These equations capture more narrow classes of behaviors than do SPQR equations, and so may necessitate more search to find a model consistent with, approximating, or appropriate to the available information, but this appears to be an advantage in light of the inherent inability of SPQR equations to express many important qualitative distinctions, however long one searches.

One might still prefer SPQR equations to parametric and piecewise-linear equations on the grounds that they are easier to analyze, in the sense that one can draw many of the most important conclusions that follow from individual SPQR equations from qualitative simulations, while many parametric and piecewise-linear equations have no closed-form solutions from which to draw the major conclusions. That is, one might view SPQR methods as trading expressive power to gain inferential efficiency. But this comparison does not represent a true tradeoff of this kind. Scientists and engineers do not restrict themselves to seeking only closed-form solutions; they also resort to numerical experiments to see if their models capture the important qualitative distinctions they observe in the world. To do this, they use their mathematical and physical knowledge to choose numerical simulations that produce reliable answers to specific qualitative questions.

We did not intend our discussion of the utility of numerical experiments to say that such computations obviate the need to reason about models. To the contrary, we followed Hamming (1962) and Truesdell (1984) in observing that setting up and accurately interpreting numerical simulations of complex systems requires a great deal of the qualitative knowledge offered by mathematics. In our view, numerical experiments correspond to observations, using a calculating machine as a sensor that detects properties of models. They cannot supplant reasoning with mathematical and physical knowledge because observations cannot be interpreted except by using knowledge of what one might be looking at or looking for. Uninformed numerical computation never substitutes for reasoning, but informed numerical calculation sometimes can.

3.2.2 Limitations of SPQR extensions

Our conclusion from the preceding observations was that SPQR does not seem to offer a better alternative than direct automation of mathematics for automating scientific and engineering reasoning. However, we recognized that no one proposes to use SPQR alone toward this end, since most current work studies various extensions to SPQR. But we did not think that such incremental extension of SPQR offers a better approach than direct automation of mathematics either. Any requirement of backward compatibility in the sense of always using concepts that refine those of SPQR enshrines the irrelevant distinctions already made in SPQR concepts. As noted earlier, the appropriate abstractions for some problems must hide the distinctions made by SPQR and introduce entirely different ones. Thus no artificial scientist or engineer will force every problem into any fixed representation, since different problems make different distinctions relevant. Relying on model representations based on those of SPQR will then be an impediment to automation, not an advantage. Direct automation of mathematical concepts will incorporate SPQR concepts and representations, of course, since they capture some important mathematical notions. An artificial scientist or engineer will, in this way, be an "extension" of SPQR, but only in a trivial sense, for it seems unlikely that SPQR concepts and representations will enter into any but a small fraction of the reasoner's abilities. We concluded that incremental extension of SPQR does not offer advantages over direct automation of mathematical knowledge.

Incremental extension of SPQR seems most attractive if one believes that SPQR captures naive or commonsense reasoning well, and that experts differ from novices and laymen mainly in possessing more knowledge and reasoning methods. We did not intend our paper to take any position on the nature of commonsense reasoning, and remain open to the possibility that SPQR aptly characterizes some forms of commonsense reasoning. But we read the psychological literature on expertise to contain substantial disagreement about the relation between naive and expert reasoning, with some psychologists, notably Carey (1985), finding that novices and experts reason with very different concepts, even though they may use some of the same words to name them. Rather than simply augmenting their commonsense concepts as they learn, experts replace their initial viewpoint with very different ways of thinking. We thus counted the view that scientific and engineering reasoning can be automated by building on SPQR as a substantial hypothesis that both requires careful justification and risks seriously impeding progress if it is wrong.

3.2.3 Advantages of direct automation

One advantage of direct automation of mathematics over incremental extension of SPQR consists of the fixed target (loosely speaking) that mathematics provides. For the purpose of automating scientific and engineering reasoning, it seems prudent to attempt to exploit the knowledge and methods that scientists and engineers have found useful, particularly the knowledge and methods that have already been formalized. Thus the researcher can use existing mathematics as a specific target for automation, and lay out a schedule (even if overly ambitious) for automating each of its theories. Simply seeking to extend SPQR incrementally runs the real risk of increasing the work necessary by not being ambitious enough. That is, without the global perspective offered by deliberately aiming to automate all of mathematical knowledge, specific problems constitute the primary impetus to incremental extensions. Successively incorporating just enough ideas to solve each specific problem may yield many special-case solutions, each of which may require just as much effort to devise as a general solution incorporating the larger theory. Even when direct automation is attempted in successive stages, at least one can choose the successive approximations to broaden the coverage, rather than risk automating numerous special cases all subsumed by the same broader approximation.

4 Reiteration

> *Now all has been heard; here is the conclusion of the matter*
> (Kohelet/Ecclesiastes 11:10, NIV)

We conclude this epilegomenon by reiterating the points of our paper.

1. Qualitative physics seeks (among other things) to automate scientific and engineering reasoning.
2. Much (if not most) work in qualitative physics has been based on the concepts, representations, and algorithms we call SPQR.
3. SPQR lacks the expressiveness and knowledge needed to automate scientific and engineering reasoning.
4. Backward compatibility with SPQR imposes high costs and probably impedes progress.
5. Mathematics provides a large vocabulary (already formalized) of qualitative concepts and results for expert and commonsense reasoning.
6. Exploiting known mathematics as rapidly as possible constitutes the most expeditious way of automating scientific and engineering reasoning.
7. Inventing new mathematics remains an opportunity for AI, but the opportunity closest to home for AI is automating the formulation and construction of models, since mathematics provides only concepts and knowledge, not automatic procedures.

References

Browder, F. E. and Mac Lane, S. The relevance of mathematics. In Steen, L. A., editor, *Mathematics Today: Twelve Informal Essays*, pages 323–350. Springer-Verlag, New York, 1978.

Carey, S. *Conceptual Change in Childhood.* MIT Press, Cambridge, MA, 1985.

COSRIMS. *The Mathematical Sciences: A Collection of Essays.* MIT Press, Cambridge, MA, 1969. Edited by the Committee on Support of Research in the Mathematical Sciences (COSRIMS) with the collaboration of George A. W. Boehm.

Doyle, J. What is rational psychology? toward a modern mental philosophy. *AI Magazine*, 4(3):50–53, 1983.

Doyle, J. Big problems for artificial intelligence. *AI Magazine*, 9(1):19–22, 1988.

Hamming, R. W. *Numerical Methods for Scientists and Engineers.* McGraw-Hill, 1962.

Jaffe, A. M. Ordering the universe. *SIAM Review*, 26(4):473–500, 1984.

Mac Lane, S. *Mathematics: Form and Function.* Springer-Verlag, New York, 1986.

Sacks, E. P. Automatic analysis of one-parameter planar ordinary differential equations by intelligent numerical simulation. *Artificial Intelligence*, 48(1):27–56, 1991.

Truesdell, C. *A First Course in Rational Continuum Mechanics*, volume 1. Academic Press, New York, 1977.

Truesdell, C. The computer: ruin of science and threat to mankind. In *An Idiot's Fugitive Essays on Science: Methods, Criticism, Training, Circumstances*, pages 594–631. Springer-Verlag, New York, 1984.

Wigner, E. P. The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics*, 13:1–14, 1960.

Yip, K. M.-K. *KAM: Automatic Planning and Interpretation of Numerical Experiments Using Geometrical Methods.* AI-TR 1163, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA, 02139, 1989.

Prolegomena

1 Introduction **1**

2 The SPQR methodology **3**
2.1 SPQR equations, behaviors, and algorithms 4
2.2 Example: tubular fluid flow 4
2.3 SPQR equations versus confluences 5
2.4 Extensions to SPQR 6
2.5 The state of the art 7

3 SPQR versus expert reasoning **9**
3.1 Experts focus on asymptotic behavior 9
3.2 Experts hypothesize and revise specific equations 12
3.3 Experts derive qualitative information by numerical analysis 14

4 Mathematics and expert reasoning **15**
4.1 Mathematical concepts serve practical modeling needs 15
4.2 Mathematical reasoning can be automated 17

5 Conclusion **19**

A Classification of SPQR examples **20**

B Classification of SPQR equations **21**

References for Prolegomena **23**

Epilegomenon

1 Our intention **27**

2 Our thesis **28**

3 Our arguments **30**
3.1 The efficacy of mathematics 30
3.1.1 The design argument 30
3.1.2 The inductive argument 31
3.1.3 The pragmatic argument 31
3.2 The advantages of mathematics 32
3.2.1 Limitations of SPQR 32
3.2.2 Limitations of SPQR extensions 35
3.2.3 Advantages of direct automation 35

4 Reiteration **36**

References for Epilegomenon **36**