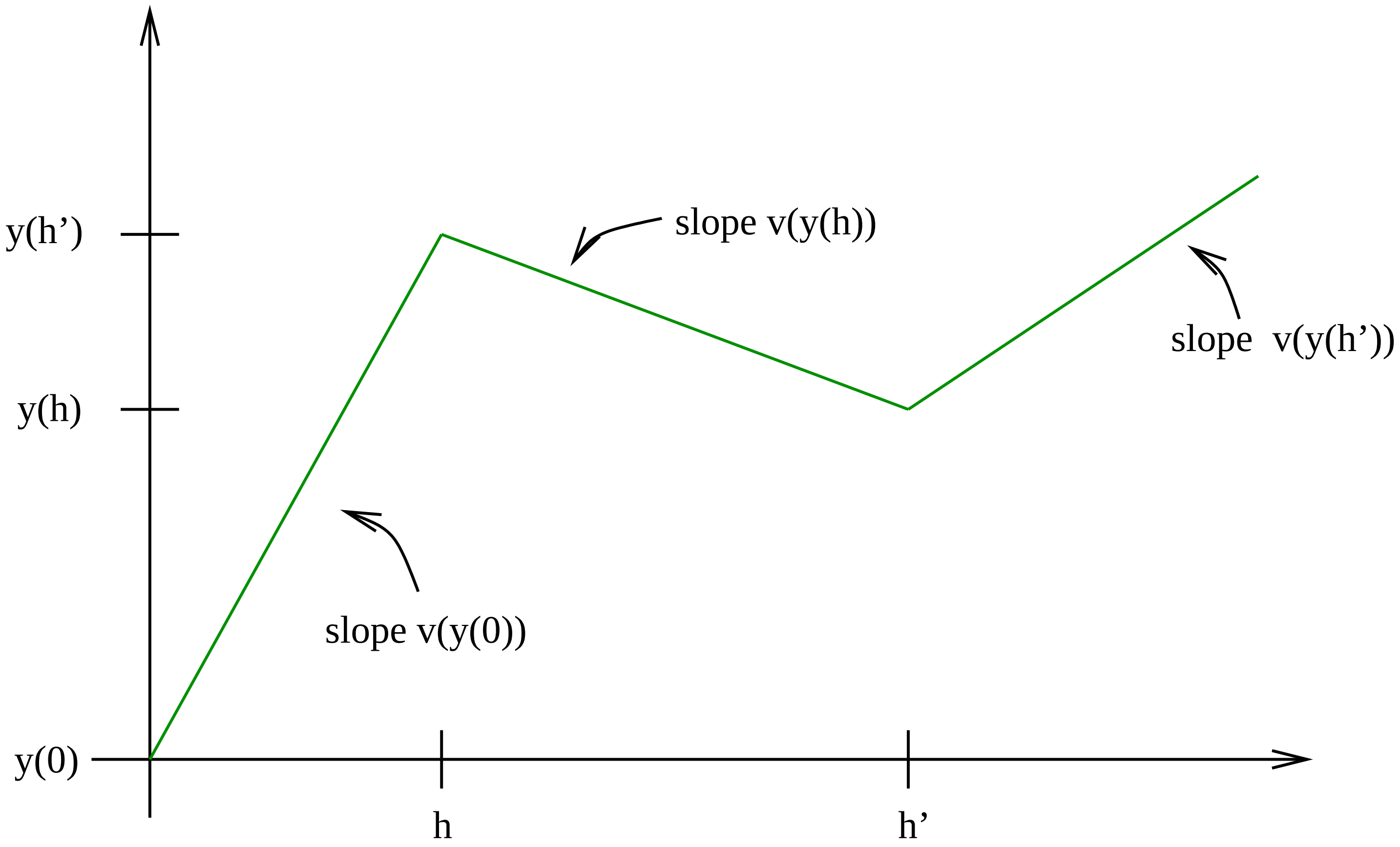
Data Types for Differential Equations

or: functional programming over mathematical structures

Asim Oomatia, Dirk Pattinson, Imperial College London

Conventional Techniques: Euler Polygons

Example. $y' = v(y), y(0) = y_0$



Taylor Series

Given. $y' = v(y), y(0) = y_0$ and a partition $(p_0, \ldots, p_k)$ with $\Delta_i = p_i - p_{i-1}$

First Order Approximations (really the mean value theorem): error $\mathcal{O}(\Delta_i)$

$$y(p_i) = y(p_{i-1}) + \Delta_i \cdot y'(\xi) \qquad (\Delta_i = p_i - p_{i-1}, \xi \in [p_i, p_{i-1}])$$

$$y(p_i) \approx y(p_{i-1}) + \Delta_i \cdot v(y(p_{i-1}))$$

Data Types. `float` :-(or intervals :-) and representation of primitive functions

Second Order Approximations (really Taylor's theorem): error $\mathcal{O}(\Delta_i^2)$

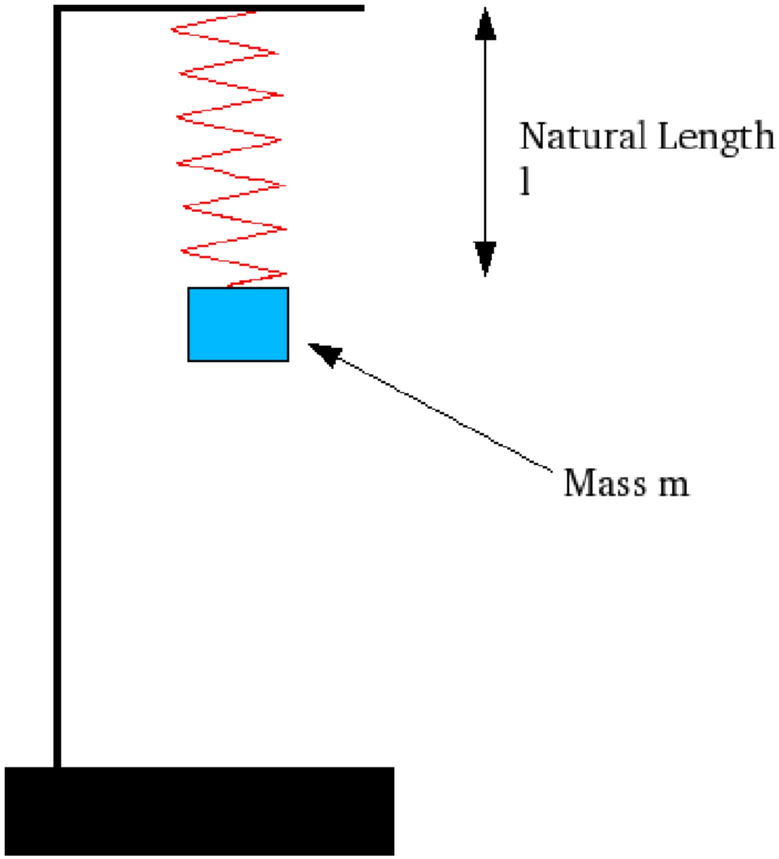
$$y(p_i) = y(p_{i-1}) + \Delta_i \cdot y'(p_i) + \frac{\Delta_i^2}{2} y''(\xi)$$

$$y(p_i) \approx y(p_{i-1}) + \Delta_i \cdot v(y(p_{i-1})) + \frac{\Delta_i^2}{2} v' \cdot v(y(p_{i-1}))$$

Data Types. Primitive functions *and their derivatives*

Modularity. Derivatives of compound functions via chain rule

The Mass on a Spring Problem



Two regimes of Evolution:

- $y \leq l$: Gravity
- $y \geq l$: Gravity plus Hooke's law

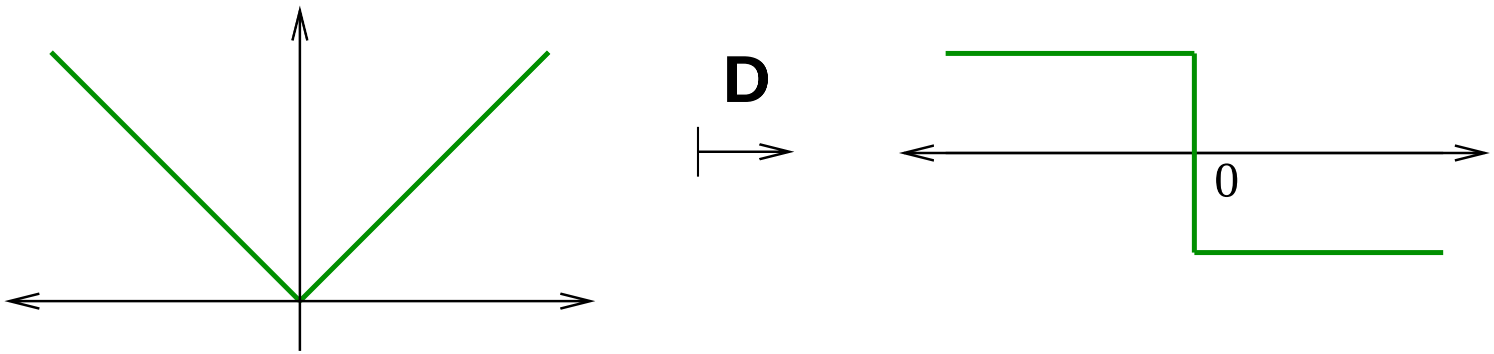
As a Differential Equation.

$$y'' = \begin{cases} g - \frac{k}{m} y & y \geq 0 \\ g & y \leq 0 \end{cases}$$

Crucially. The (function defining the) differential equation is *Lipschitz*, but *not* C^1.

But Derivatives can be Sandwiched

Example. Absolute value $f(x) = |x|$



Example. $f(x, y) = \max(x, y)$

$$\mathsf{D}(f)(x,y) = \begin{cases} (1,0) & x > y \\ [0,1] \times [0,1] & x = y \\ (0,1) & x < y \end{cases}$$

New Data Types. Derivatives (and, *a posteriori*, functions) are *interval valued*.

- what is the *mathematical theory* of interval derivatives?
- how about *convergence speed* and relation to conventional methods?

Main Ideas

Goal. Use of *higher order* methods $\rightsquigarrow$ *much* improved convergence behaviour

Main Idea.

- trade the *real line* $\mathbb{R}$ for the *interval domain* $\mathbf{I}\mathbb{R} = \{[a, b] \mid a, b \in \mathbb{R}, a \leq b\}$
- trade *classical* for *domain theoretic* (Clarke's) derivatives

Digression. Domain Theoretic Differential Calculus

- Taylor's Theorem
- the Chain Rule

Buy one, get one free – and some more goodies

- two orders (of differentiability) for the price of one
- plus: guaranteed (much improved) error bounds
- plus: "proof-by-calculation"-style existence assertions

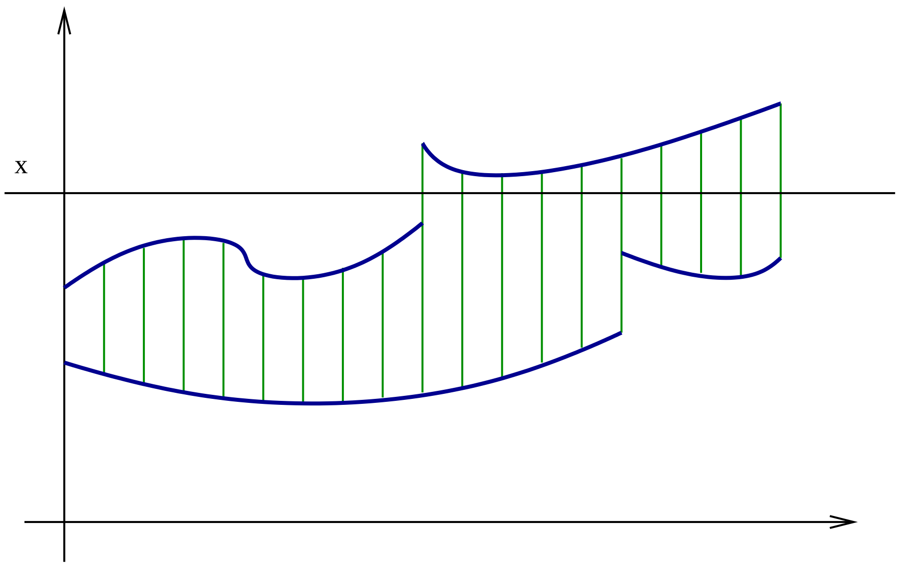
Domain Theoretic Differential Calculus

Outset. Consider functions $f = [\underline{f}, \overline{f}] : O \to \mathbf{I}\mathbb{R}$ where

- $O \subseteq \mathbb{R}^n$ is (usually) open
- $(\mathbf{I}\mathbb{R}, \sqsubseteq)$ is the *interval domain* of compact intervals ordered by reverse inclusion

Topology. Consider $\mathbf{I}\mathbb{R}$ as equipped with the *Scott topology* so that

$$f \text{ continuous } \iff \underline{f}, \overline{f} \text{ upper (lower) semi continuous}$$



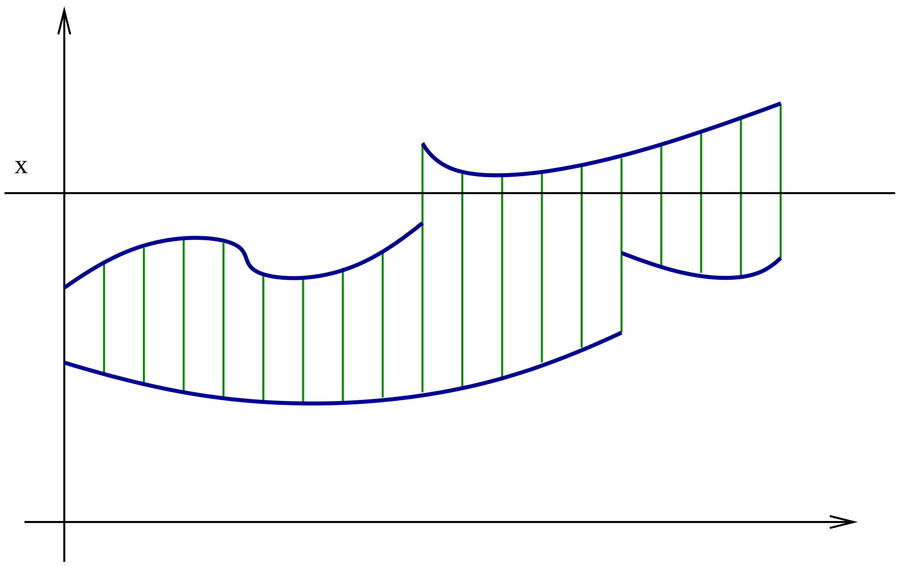
Domain Theoretic Differential Calculus

Outset. Consider functions $f = [\underline{f}, \overline{f}] : O \to \mathbf{I}\mathbb{R}$ where

- $O \subseteq \mathbb{R}^n$ is (usually) open
- $(\mathbf{I}\mathbb{R}, \sqsubseteq)$ is the *interval domain* of compact intervals ordered by reverse inclusion

Topology. Consider $\mathbf{I}\mathbb{R}$ as equipped with the *Scott topology* so that

$$f \text{ continuous} \iff \underline{f}, \overline{f} \text{ upper (lower) semi continuous}$$



Example. Intermediate Value Theorem.

If $f : [a,b] \to \mathbf{I}\mathbb{R}$ is continuous and

$$x \in f(a) \sqcap f(b)$$

then $x \in f(t)$ for some $t \in [a,b]$.

The Domain Theoretic Derivative

Informal Definition.

Suppose $f : O \to \mathbb{R}$ and $a \subseteq O \subseteq \mathbb{R}$ are open.

$$\mathsf{D}_a(f) = \bigsqcap_{x \neq y \in a} \frac{f(y) - f(x)}{y - x} \qquad\qquad \mathsf{D}(f)(x) = \bigsqcup_{x \in a} \mathsf{D}_a(x)$$

Idea. $\mathsf{D}(f)(x)$ contains all slopes that appear around x.

Formal Definition.

Suppose $f : O \to \mathbb{R}$ and $O \subseteq \mathbb{R}^n$ is open.

Then f has a Lipschitz constant $b \in \mathbf{I}\mathbb{R}^n$ around the open set $a \subseteq O$, written

$$f \in \delta(a, b) \text{ iff } \forall x, y \in a^o . b(x - y) \sqsubseteq f(x) - f(y)$$

and the *Domain Theoretic Derivative* is the sup of all Lipschitz constants

$$\mathsf{D}(f) = \bigsqcup_{f \in \delta(a,b)} a \searrow b \qquad \text{where } a \searrow b(x) = b \text{ if } x \in a \text{ and } (-\infty, \infty) \text{ otherwise}$$

Some Properties

Existence and Continuity. $\mathsf{D}(f)$ is Scott continuous and $\mathsf{D}(f) = f'$ if $f \in \mathcal{C}^1$.

Relationship to Clarke's Gradient

- $\mathsf{D}(f)$ is the "smallest rectangular approximation" of Clarke's gradient
- equality if we work with compact sets rather than products of intervals

Examples.

- $f(x) = |x| \rightsquigarrow \mathsf{D}(f)(0) = [-1, 1]$ as expected
- $f(x) = \sqrt{|x|} \rightsquigarrow \mathsf{D}(f)(0) = \perp$

More generally. $\mathsf{D}(f)(x) \neq \perp$ whenever f is Lipschitz in a neighbourhood of x.

$\rightsquigarrow$ locally Lipschitz vector fields (also guarantees uniqueness)

The Chain Rule

Notation

- If $f = (f_1, \ldots, f_k) : O \to \mathbb{R}^k$ then $\mathsf{D}(f) = (\mathsf{D}(f_1), \ldots, \mathsf{D}(f_k))$
- If $f : O \to \mathbf{I}\mathbb{R}^m$ and $g : P \to \mathbf{I}\mathbb{R}^n$ for $O \subseteq \mathbb{R}^n$ with $g(x) \subseteq O$

$$f \circ g(x) = \bigsqcap_{y \in g(x)} f(y)$$

The Chain Rule.

If $f : O \to P$ and $g : P \to \mathbb{R}^k$ where $O \subseteq \mathbb{R}^n$ and $P \subseteq \mathbb{R}^m$ are open, then

$$\mathsf{D}(f \circ g) \sqsupseteq \mathsf{D}(f) \circ g \cdot \mathsf{D}(g)$$

Equality may fail: Suppose $f(x) = |x|$ and $g(x) = \max(x, 0)$. Then

$$\mathsf{D}(f \circ g)(0) = \mathsf{D}(g)(0) = [0, 1]$$

$$\mathsf{D}(f) \circ g \cdot \mathsf{D}(g)(0) = [-1, 1] \cdot [0, 1] = [-1, 1]$$

("*smoothing out*", also e.g. for $f(x) = x^2$ and $g(x) = |x|$)

The Mean Value Theorem.

Suppose $f : [a, b] \longrightarrow \mathbb{R}$ is *continuous* and $x \in [a, b]$. Then

$$f(x) \in f(a) + (x - a)\mathsf{D}(f)(\xi)$$

for some $\xi \in [a, x]$.

Taylor's Theorem. Suppose $f : [a, b] \in \mathcal{C}^n[a, b]$. Then

$$f(x) \in f(a)+(x-a)f'(a)+\cdots+\frac{(x-a)^n}{n!}f^{(n)}(a)+\frac{(x-a)^{n+1}}{(n+1)!}\mathsf{D}(f^{(n)})(\xi)$$

for some $\xi \in [a, x]$.

Remark. Taylor's Theorem holds *even without the assumption* $f \in \mathcal{C}^n$ if we consider derivatives of functions $f : O \longrightarrow \mathbf{I}\mathbb{R}$.

Back to Differential Equations . . .

Problem. Find a solution of the initial value problem

$$y' = v(y), y(t_0) = a$$

where $O \subseteq \mathbb{R}^n$ is open, $v : O \to \mathbb{R}^n$ is locally Lipschitz and $a \in \mathbb{R}^n$.

Our Setup. We start with two functions

- $u : \mathbf{I}O \to \mathbf{I}\mathbb{R}^n$ that *extends* v, i.e. $u(x) \sqsubseteq v(x)$ for all $x \in O$.
- $u' : \mathbf{I}O \to \mathbf{I}\mathbb{R}^{n\times n}$ that *extends* $\mathsf{D}(v)$, i.e. $u'(x) \sqsubseteq \mathsf{D}(v)(x)$ for all $x \in O$.

Existence Theorem. Suppose that $\alpha, E \in \mathbf{I}\mathbb{R}^n$ and

$$u(\alpha + [0, h] \cdot E) \subseteq E$$

Then, for all $a \in \alpha$, the problem has a solution $z : [t_0, t_0 + h] \to \mathbb{R}^n$ that satisfies

$$z(t) \in \alpha + (t - t_0) \cdot u(\alpha + [0, h] \cdot E) \subseteq \alpha + [0, h] \cdot E$$

Algorithmic Digression

Given. $\alpha \in \mathbf{I}\mathbb{R}^n$ so that the solution z of the IVP satisfies $z(t') \in \alpha$

- Pick a step size h
- Guess a *valid enclosure* E s.t. $u(\alpha + [0, h] \cdot E) \subseteq E$
- Put $y(t) = \alpha + (t' - t)u(\alpha + [0, h] \cdot E)$ for $t \in [t', t' + h]$

Then the solution z will satisfy $z(t) \in y(t)$ for $t \in [t', t' + h]$.

Proof by Calculation. Repeating this process for $t = t_0, t_0 + h_0, \ldots, t_0 + h_k$ gives

- a *proof of existence* of a solution in $[t_0, t_0 + h_k]$
- guaranteed (first order) *bounds on the values* of a solution.

Parameters to fiddle. Enclosure E and step size h.

Second Order Steps

Enclosure Theorem (2nd Order Version) Suppose E is a valid enclosure at α for time h, i.e.

$$u(\alpha + [0,h] \cdot E) \subseteq E$$

Then every solution $z : [t_0, t_0 + h]$ of $y' = v(y), y(t_0) = a \in \alpha$ satisfies

$$z(t) \in \alpha + (t - t_0)u(\alpha) + \frac{(t-t_0)^2}{2} u' \cdot u(\alpha + [0,h] \cdot E)$$

Proof. Chain Rule and Taylor's Theorem.

Compare with *first order approximation*:

$$z(t) \in \alpha + (t - t_0)u(\alpha + [0,h] \cdot E)$$

- overestimation is *quadratic*
- significant gain *in practice*, but *not in theory*

Comparision

Under the Carpet: Intricacies of Guessing, Monotonicity and Assembly of Solutions

Informal Notation. Let $y_1(\mathcal{P})$ (resp. $y_2(\mathcal{P})$) denote the gluing together of k first (second) order steps.

Montonicity. If $\mathcal{Q}$ is a refinement of $\mathcal{Q}$, then $y_i(\mathcal{P}) \sqsubseteq y_i(\mathcal{Q})$, $i = 1, 2$.

(guessing for a finer partition improves guessing for a coarser partition)

Convergence. If $(\mathcal{P}_k)$ is a sequence of partitions that refines exponentially fast, then $y_i(\mathcal{P}_k) \longrightarrow z$ exponentially fast, where $z(t_0) = a$, $z' = v(z)$.

Striking Fact. Cannot *prove* that 2nd order converges faster.

- if $\mathsf{D}(v)$ is bounded by L, then nL is a Lipschitz constant for v
- first oder convergence also depends on Lipschitz constant, and
- 2nd order convergence speed is a *global* estimate

Types, Concretely

Vector Fields: symbolic representation

```
data Term = Plus Term Term |  ... | Max  Term Term | Abs Terr
```

(Dependent Types would be nice to have)

Evaluation.

```
eval :: Term -> Vector -> Interval
eval (Plus s t) v = iadd (eval s v) (eval t v)
eval (Max s t)  v = imax (eval s v) (eval t v)
```

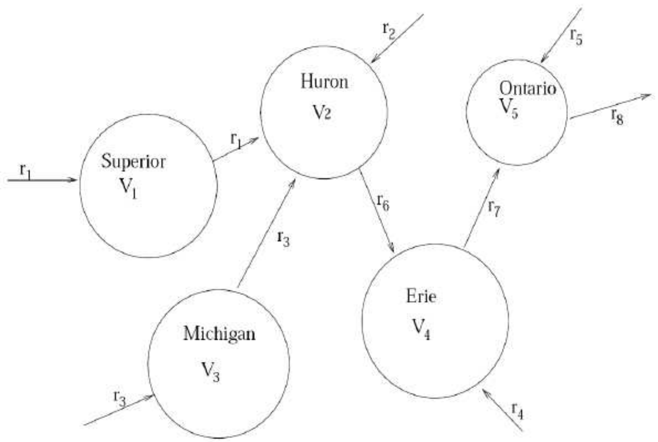
Differentiation. Symbolically, e.g.

```
deriv :: Term -> Vector -> Vector
```

(using the chain rule)

Experimental Results: Differentiable Vector Fields

Lake Pollution Model (restricted to Lake Huron, Michigan and Superiour)



Stepsize	Width		Run Time		Markup
	1^{st} Ord	2^{nd} Ord	1^{st} Ord	2^{nd} Ord	
0.1	1.36	0.0684	0.57	0.45	200
0.01	0.0367	0.0000184	9.2	4.63	2000
0.001	0.00301	0.0000002	98.9	57.4	15060

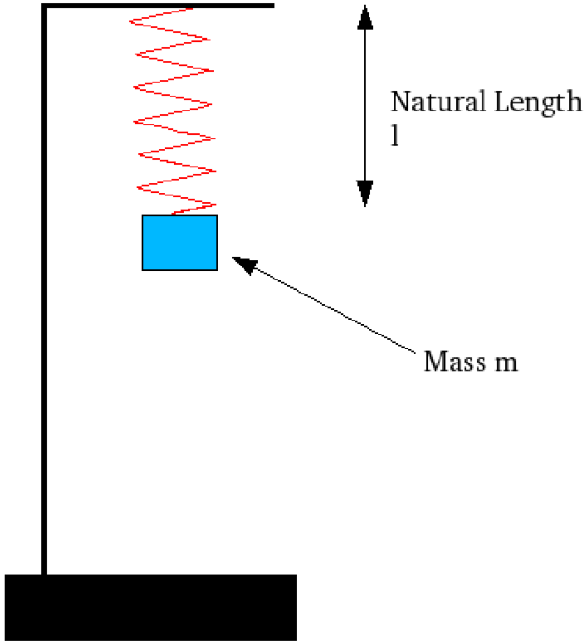
Accuracy at $t = 10$

Decline of Pollution in the network of Great Lakes in North America over time

$$v(y_1, y_2, y_3) = \left(-\frac{r_1}{V_1}y_1, \frac{r_1}{V_1}y_1 + \frac{r_3}{V_3}y_3 - \frac{r_1 + r_2 + r_3}{V_2}y_2, -\frac{r_3}{V_3}y_3\right)$$

Note. The second order method is not only *better* but also *faster*!

Mass on a Spring



Stepsize	Width		Run Time		Markup
	1^{st} Ord	2^{nd} Ord	1^{st} Ord	2^{nd} Ord	
0.01	1105	42	32.3	2	26.3
0.001	212	0.38	2610	61	557

Accuracy at $t = 8$

$$v(y_1, y_2) = (y_2, -\max(-g, \frac{k}{m}x - g))$$

Initial values: $y_1 = y_2 = 0$

Some Conclusions

Implementation is Important.

- also uncovers mathematically interesting questions
- reveals unexpected phenomena
- identifies shortcomings of the theory

Conventional Techniques.

- require *existence* of classical derivatives
- no guaranted bounds

Calculus in Domain Theory.

- uses *generalised* derivatives that *always exist*
- automatic *existence proof* as a by-product