Recitation 6: Differential Invariants and Differential Cuts
15-424/15-624/15-824 Logical Foundations of Cyber-Physical Systems

Notes by Brandon Bohrer.
Edits by Yong Kiam Tan, (later) Katherine Cordwell (kcordwel@cs.cmu.edu), and Aditi Kabra

1 Motivation and Learning Objectives

We shall start the recitation by reviewing the ODE axioms of dL that we have seen in class: differential invariants, differential cuts. By piecing these axioms together with the other basic ODE axioms, we can derive powerful *proof rules* for reasoning about ODEs.

Using these proof rules, we will take another look at the time-triggered ping-pong model from last week. Specifically, we will focus on proving properties about the ODEs used to model the motion of the ball. The aforementioned proof rules can be used to prove more advanced properties of these ODEs.

2 Axioms and Proof Rules for Differential Equations

The differential equations axioms for dL can be roughly characterized into two groups. In the first group, we have the basic "helper" axioms that allow us to syntactically extract information from an ODE. The latter group of axioms are the ones that require proof insights from the user.

2.1 ODE Helper Axioms

The differential effect and differential weakening axioms belong to the first group of helper axioms:

$$\text{(DE)}\quad [\{x' = f(x)\,\&\,Q\}]P \leftrightarrow [\{x' = f(x)\,\&\,Q\}][x' := f(x)]P$$

$$\text{(DW)}\quad [\{x' = f(x)\,\&\,Q\}]P \leftrightarrow [\{x' = f(x)\,\&\,Q\}](Q \rightarrow P)$$

The differential effect axiom (DE) extracts information about the differential equations $x' = f(x)$. It internalizes that along any solution to the differential equation, the differential variables x' must take on the correct values, i.e., the values of the RHS $f(x)$. In order to prove a postcondition P of an ODE, the RHS of axiom DE says that it suffices to prove the postcondition $[x' := f(x)]P$ instead. The assignment $x' := f(x)$ allows us replace all of the differential variables in P with their respective right-hand sides.

Similarly, the differential weakening axiom (DW) extracts information about the domain constraint Q. Just like the differential variables must take on the correct values, the domain constraint Q must be satisfied along any solution to the differential equation. Therefore, in order to prove a postcondition P of an ODE, the RHS of axiom DW says that it suffices to

prove the postcondition $Q \to P$ instead, where we are allowed to assume that the domain constraint Q is still true at the end of a solution.

It is straightforward to derive proof rules for these two axioms using G. Notice in both cases that the use of G forces us to drop the assumptions Γ, Δ in the context:

$$\text{(dE)} \quad \frac{Q \vdash [x' := f(x)]P}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}$$

$$\text{(dW)} \quad \frac{Q \vdash P}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}$$

Exercise 1:
Work through the derivation of both proof rules. Make sure you see where the additional context has to be dropped in both rules.
Answer:

$$\cfrac{\cfrac{Q \vdash P}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}](Q \to P), \Delta}\,{\scriptstyle \text{G},\to\text{R}}}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}\,{\scriptstyle \text{DW}}$$

$$\cfrac{\cfrac{Q \vdash [x' := f(x)]P}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}][x' := f(x)]P, \Delta}\,{\scriptstyle \text{dW}}}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}\,{\scriptstyle \text{DE}}$$

As usual, constant assumptions in the context can be kept and KeYmaera X will do this for you automatically. By themselves, however, the dE and dW proof rules are not very useful. It is the next group of axioms for ODEs which really allow us to start proving interesting properties. This group includes the differential invariant and differential cut axioms, listed here for easy reference:

$$\text{(DI)} \quad \big([\{x' = f(x) \,\&\, Q\}]P \leftrightarrow [?Q]P\big) \leftarrow (Q \to [\{x' = f(x) \,\&\, Q\}](P)')$$

$$\text{(DC)} \quad \big([\{x' = f(x) \,\&\, Q\}]P \leftrightarrow [\{x' = f(x) \,\&\, Q \wedge C\}]P\big) \leftarrow [\{x' = f(x) \,\&\, Q\}]C$$

2.2 Differential Invariants

The differential invariants axiom is the core workhorse axiom for proving properties of ODEs. In particular, it gives us a link between postcondition P and its differential $(P)'$. This is the crucial insight that we discussed last week: instead of working directly with the solution, we shall instead work with their derivatives (and thus, the ODEs) directly.

The notation $(P)'$ can be understood as a generalization of the differential of terms $(e)'$ to a differential on formulas. We can think of it as a shorthand that we are designing specifically for use in the differential invariants axiom/proof rule. It is defined inductively, with the following base cases for the atomic comparison formulas:

$$
\begin{aligned}
(e = k)' &\equiv (e)' = (k)' \\
(e \leq k)' &\equiv (e)' \leq (k)' \qquad \text{(accordingly for } \geq) \\
(e < k)' &\equiv (e)' \leq (k)' \qquad \text{(accordingly for } >) \\
(e \neq k)' &\equiv (e)' = (k)'
\end{aligned}
$$

Note: The differential of $(e < k)'$ could also be defined as $(e)' < (k)'$. This would also be sound, but just more restrictive than the one given above.

Let us look specifically at the $e = k$ case for some intuition:

(DI$_=$) $\big([\{x' = f(x)\}]e = k \leftarrow e = k\big) \leftarrow ([\{x' = f(x)\}](e)' = (k)')$

In this equational case, the axiom says that in order to prove that $e = k$ is true along all solutions to the ODE, it suffices to show that $e = k$ is true initially and that $(e)' = (k)'$ is true along all solutions to the ODE. Why might this be the case? The answer comes from the crucial differential lemma from last week.

Remember that differentials $(e)'$ have the same value as the derivative of e along solutions of an ODE. Therefore, if the values of terms e and k start off equal, and the value of their derivatives stay equal along the solution, then the values e and k also stay equal along the solution, i.e., $e = k$ is true along the solution.

Exercise 2:

Where should we add back the domain constraints in the DI$_=$ axiom, and why (intuitively)?

Answer: By the original DI axiom, the equational case with domain constraints is as follows:

(DI$_=$) $\big([\{x' = f(x) \,\&\, Q\}]e = k \leftrightarrow [?Q]P\big) \leftarrow (Q \rightarrow [\{x' = f(x) \,\&\, Q\}](e)' = (k)')$

It would be sound to remove the $Q \rightarrow$. There are some cases where including the $Q \rightarrow$ is helpful. For example, try to prove $x = 1 \rightarrow [x' = f(x) \;\&\; x \neq 1]x \geq 1$ with both versions of the axiom and with the dI proof rule, which is actually derived from the version of the axiom that does not include the $Q \rightarrow$.

Recall that the domain constraint Q of an ODE must be obeyed at *all times*, including at the start. Thus, in an initial state where Q is false, the formula $[\{x' = f(x) \,\&\, Q\}]e = k$ is vacuously true. The DI$_=$ axiom tells us exactly that, because both $[?Q]P$ and $Q \rightarrow [\{x' = f(x) \,\&\, Q\}](e)' = (k)'$ would be vacuously true in such a state.

More interestingly, we now only need to prove $[\{x' = f(x) \,\&\, Q\}](e)' = (k)'$, rather than $[\{x' = f(x)\}](e)' = (k)'$. Recall additionally, that the axiom DW allows us to assume Q when proving the postcondition of an ODE. Thus, we can now prove $(e)' = (k)'$ while assuming the domain constraint Q. This intuition can be more easily seen from the proof rule:

(dI$_=$) $$\frac{Q \vdash [x' := f(x)](e)' = (k)'}{\Gamma, e = k \vdash [\{x' = f(x) \,\&\, Q\}]e = k, \Delta}$$

Exercise 3:

Derive this proof rule from the DI$_=$ axiom.

Answer: This derives using dE which in turn derived from dW.

As an aside: why do we want to write down these proof rules when we could have just derived them from the axioms? Proof rules provide a useful summary of the standard way in which we would put the axioms together to prove a desired postcondition. They are well suited, e.g., for use as top-level tactics in KeYmaera X, because that is what you would want to work with rather than applying the axioms step by step every single time you want to prove that something is invariant for the ODEs.

The intuition behind the other base cases is similar to the $e = k$ case. In order to prove that $e \geq k$ (resp. $e > k$) is true along a solution of the ODE, we will require that $e \geq k$ (resp. $e > k$) initially, and that $(e)' \geq (k)'$ along that solution, or in other words, the value of the derivative of e is always greater or equal to that of k along the solutions.

For $e \neq k$, it is initially slightly surprising that we instead need to show that $(e)' = (k)'$ along the ODE rather than $(e)' \neq (k)'$. This surprise should clear up, however, once we realize that checking $(e)' \neq (k)'$ is insufficient to ensure that $e \neq k$ stays true along the ODE. An an example, consider the following sequent:

$$x \neq y \vdash [\{x' = 1, y' = -1\}]x \neq y$$

Exercise 4:
Is this sequent valid? Why or why not?
Answer: It is clearly not valid: consider an initial state where $x < y$, then since the ODEs increase x while decreasing y, their values should eventually meet somewhere along the solution to the ODE.

If we had defined $(e \neq k)' \equiv (e)' \neq (k)'$, however, we would have easily proved the above property (unsoundly):

$$\cfrac{\cfrac{\cfrac{*}{\vdash 1 \neq -1}}{\vdash [x' := 1][y' := -1]x' \neq y'}[':=]}{x \neq y \vdash [\{x' = 1, y' = -1\}]x \neq y}\,??$$

Finally, the $(\cdot)'$ operator can be extended to conjunctive and disjunctive formulas inductively. This extension can also be thought of as a mnemonic device, since we already saw in class that both the $\wedge$ and $\vee$ cases can be derived from the base cases using the other proof rules of dL.

$$(P \wedge Q)' \equiv (P)' \wedge (Q)'$$
$$(P \vee Q)' \equiv (P)' \wedge (Q)'$$

Like the $\neq$ case, the $\vee$ case is slightly surprising: it requires that we prove $(P)' \wedge (Q)'$ rather than $(P)' \vee (Q)'$ as we might expect if we simply extended the $(\cdot)'$ operator naively. However, there is good reason for this. If Q is true initially (but P is false), then knowing that $(P)'$ holds along the solutions to an ODE doesn't tell us anything about whether Q

stays true (and in fact P could stay false). For example, we don't want to be able to prove $(4 > 5 \lor x = 1) \vdash [\{x' = 2\}](4 > 5 \lor x = 1)$.

Another way of understanding this difference is to consider the $e \neq k$ case we saw above. In real arithmetic, the formula $e \neq k$ can be re-written equivalently as $e > k \lor e < k$. If we simply defined the differential of a disjunction to be $(P)' \lor (Q)'$, then the differential of $e > k \lor e < k$ case would be $(e)' \geq (k)' \lor (e)' \leq (k)'$ which is a valid formula in real arithmetic. So we have just proved that $e \neq k$ is an invariant of any ODE we ever wanted simply by rewriting it into another form. In contrast, the correct definition yields $(e)' \geq (k)' \land (e)' \leq (k)'$, which is equivalent to $(e)' = (k)'$, as expected. To summarize, the proof rule for differential invariants is:

$$\text{(dI)} \quad \frac{Q \vdash [x' := f(x)](P)'}{\Gamma, P \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}$$

Exercise 5:
Derive this proof rule (similar to the derivation of $\text{dI}_=$). The answer is in the textbook.

2.3 Differential Cuts

The differential cuts axiom (DC) is very much like the usual cut rule from sequent calculus, except it allows us to cut and assume new formulas in the domain constraint of a differential equation rather than the antecedents. The proof rule for differential cuts is a straight forward rephrasing of the DC axiom with two premises:

$$\text{(dC)} \quad \frac{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]C, \Delta \qquad \Gamma \vdash [\{x' = f(x) \,\&\, Q \land C\}]P, \Delta}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}$$

Exercise 6:
Convince yourself that this rule derives from DC.

The right premise of dC allows us to additionally assume the cut C in the domain constraint when proving postcondition P. In particular, if we subsequently use dI, in the right premise, we now get to assume $Q \land C$ when proving $(P)'$:

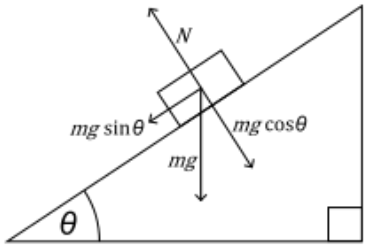
$$\text{dC}\frac{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]C, \Delta \qquad \text{dI}\dfrac{Q \land C \vdash [x' := f(x)](P)'}{\Gamma \vdash [\{x' = f(x) \,\&\, Q \land C\}]P, \Delta}}{\Gamma \vdash [\{x' = f(x) \,\&\, Q\}]P, \Delta}$$

As an aside, whereas the usual cut rule in logic can be removed by a cut elimination theorem, the dC principle cannot be eliminated without losing the ability to prove some ODE properties in dL. In other words, if we restricted ourselves to using dI without ever using dC, we would not be able to prove some true properties of ODEs that we could have with dC. We will see more about the proof theory of the ODE axioms in dL later in the course.

3 Practice

Let's practice! We'll explore differential equations in a new setting: boxes sliding down slopes. The properties we will prove about this example are relatively straightforward consequences from physics. Our main objective here is to gain some routine practice with modeling systems and doing proofs with differential invariants and differential cuts.

The first model that we shall explore is a box sliding down a frictionless slope. This is not a physics course, so we will just look at Wikipedia for a useful illustration of this situation:



From the picture, we see that the box is being driven parallel down the slope by a force $mg\sin(\theta)$. For our purposes, we shall work with accelerations and so we will assume that the box has unit mass, i.e., with $m = 1$.

To set us up for the question we will be asking next, we shall work in Cartesian coordinates i.e., in the xy-plane. In that case, we will need to describe the motion of the box along both coordinate axes.

3.1 One Slippery Slope

The coordinates simply change according to their respective components of the velocities:

$$x' = v_x, y' = v_y$$

However, since the box is *accelerating*, we will also need to model these velocities changing as it slides down the slope. Using basic trigonometry, we have:

$$v_x' = g\sin(\theta)\cos(\theta), v_y' = -g\sin^2(\theta)$$

This all makes sense, *except* the right hand sides of these ODEs are not quite polynomials so we cannot actually write these down in dL. To fix this issue, we will have to rewrite these trigonometric functions away.

Exercise 7:
How?
Answer: For this purpose, we shall introduce two new variables w, h which represent the width, height of the slope respectively. Since we will actually only be concerned with the steepness of the slope rather than its precise length we can, without loss of generality, assume

that the length of the diagonal $w^2 + h^2 = 1$. Thus, we can write $\sin(\theta) = h, \cos(\theta) = w$ and use this to rewrite the ODEs for velocity as follows:

$$v_x' = gwh, v_y' = -gh^2$$

Now that we have set up a system of ODEs, let us see whether we can prove interesting properties about it. Recall that the box was accelerating down the slope at rate $g\sin(\theta) = gh$. If its initial velocity was 0, then the distance it moves after time t should be given by $\frac{1}{2}ght^2$.

To help us write this down formally, it is useful to start with some abbreviations. First, the assumptions on the constants:

$$\Gamma \stackrel{\text{def}}{\equiv} g > 0, h^2 + w^2 = 1, h > 0, w > 0$$

Next, the system of ODEs with an additional clock equation $t' = 1$:

$$\alpha \stackrel{\text{def}}{\equiv} \{x' = v_x, y' = v_y, v_x' = gwh, v_y' = -gh^2, t' = 1\}$$

Finally, some assumptions on the initial values of the variables:

$$Init \stackrel{\text{def}}{\equiv} t = 0, x = x_0, y = y_0, v_x = 0, v_y = 0$$

This is the sequent we want to prove valid:

$$\Gamma, Init \vdash [\alpha](x - x_0)^2 + (y - y_0)^2 = \left(\frac{1}{2}ght^2\right)^2$$

In order to avoid writing down square roots we have written down the *squared* distance in the postcondition. This results in a fourth power of t appearing on the right.

There are several ways we can prove this sequent. First, the system of ODEs that we have written down is actually solvable, so it is possible to simply solve and ask QE. Second, we could try a direct proof using dI,dC which is possible, but would not be very pleasant.

Exercise 8:
How else could we prove this sequent?
Answer: A third option, which is the approach we will try next, is to instead prove a more straightforward postcondition that implies what we want. Recall that we have already factored the velocity into the x and y directions. We could simply give the closed form expression for the positions moved in both of these directions. For the horizontal x direction, we shall prove the following sequent:

$$\Gamma, Init \vdash [\alpha]x - x_0 = \frac{1}{2}gwht^2 \tag{1}$$

For space, abbreviate $x - x_0 = \frac{1}{2}gwht^2$ by P.
Let us try a straightforward dI:

$$\text{cut}\dfrac{\text{WR}\dfrac{\mathbb{R}\dfrac{*}{\Gamma, Init \vdash P}}{\Gamma, Init \vdash P, [\alpha]P} \qquad \text{dI}\dfrac{[':=]\dfrac{\vdash v_x = \frac{1}{2}gwh(2t)}{\vdash [x' := v_x][t' := 1]x' = \frac{1}{2}gh^2(2tt')}}{\Gamma, Init, P \vdash [\alpha]P}}{\Gamma, Init \vdash [\alpha]P}$$

Note: Here we explicitly cut in $x - x_0 = \frac{1}{2}gwht^2$ to our assumptions. In the rest of these notes we will leave such cuts implicit (for space considerations and also because they're not hard to prove). KeYmaera X will do such a cut automatically. However, it is important to understand the purpose of this propositional cut—it's to double check that the invariant we're trying to prove is true initially—and these cuts should be included e.g. in homework assignments.

As we have seen several times already, the proof fails because we do not have enough information about v_x. We do get a hint, however, that we should first try a differential cut of $v_x = \frac{1}{2}gwh(2t)$. This cut proves fine:

$$\text{dI}\dfrac{[':=]\dfrac{\mathbb{R}\dfrac{*}{\vdash gwh = gwh}}{\vdash [v_x' := gwh][t' := 1]v_x' = gwht'}}{\Gamma, Init \vdash [\alpha]v_x = \frac{1}{2}gwh(2t)}$$

It allows us to complete our earlier proof by first using a dC:

$$\text{dC}\dfrac{\text{dI}\dfrac{[':=]\dfrac{\mathbb{R}\dfrac{*}{v_x = \frac{1}{2}gwh(2t) \vdash v_x = \frac{1}{2}gwh(2t)}}{v_x = \frac{1}{2}gwh(2t) \vdash [x' := v_x][t' := 1]x' = \frac{1}{2}gwh(2tt')}}{\Gamma, Init \vdash [\{\alpha \,\&\, v_x = \frac{1}{2}gwh(2t)\}]x - x_0 = \frac{1}{2}gwht^2}}{\Gamma, Init \vdash [\alpha]x - x_0 = \frac{1}{2}gwht^2}$$

The vertical y direction can be proved similarly, i.e., this sequent is also valid:

$$\Gamma, Init \vdash [\alpha]y - y_0 = \frac{1}{2}gh^2t^2 \tag{2}$$

Using Equations 1 and 2, we can now prove the Euclidean distance property that we wanted using an M[·] step. The M[·] step works because if the equations $y - y_0 = -\frac{1}{2}gh^2t^2$ and $x - x_0 = \frac{1}{2}gwht^2$ are true, then using the assumption $w^2 + h^2 = 1$ we have:

$$\begin{aligned}(y - y_0)^2 + (x - x_0)^2 &= \left(\frac{1}{2}gh^2t^2\right)^2 + \left(\frac{1}{2}gwht^2\right)^2 \\ &= \left(\frac{1}{2}ght^2\right)^2 (h^2 + w^2) \\ &= \left(\frac{1}{2}ght^2\right)^2\end{aligned}$$

$$\text{M[·]}\dfrac{[]\wedge,\wedge\text{R}\dfrac{1 \qquad 2}{\Gamma, Init \vdash [\alpha]x - x_0 = \frac{1}{2}gwht^2 \qquad \Gamma, Init \vdash [\alpha]y - y_0 = -\frac{1}{2}gh^2t^2}{\Gamma, Init \vdash [\alpha](x - x_0 = \frac{1}{2}gwht^2 \wedge y - y_0 = -\frac{1}{2}gh^2t^2)}}{\Gamma, Init \vdash [\alpha](x - x_0)^2 + (y - y_0)^2 = (\frac{1}{2}ght^2)^2}$$

Exercise 9:
There is a fourth related option. We can make use of the fact that we already know the box is sliding down the slope. How?
Answer: Since the box is sliding down the slope, the following will also be an invariant:

$$y_0 - y = \frac{h}{w}(x - x_0) \tag{3}$$

This will, in turn, require us to prove the following invariant on velocities which follows easily by dI:

$$-v_y = \frac{h}{w}v_x$$

This approach is somewhat more satisfying because it gives us an actual invariant about the motion of the box that is independent of time. Using Equations 3 and 1, it is also possible to deduce the distance moved by the box.

The main takeaway messages here is that directly attempting to use dI may not always be the best option. Rephrasing the question could make it easier to prove.

3.2 Two Slippery Slopes

Suppose you were in a competition where you were asked to build a slope so that the boxes slide down and hit the floor as fast as possible. From ordinary physical intuition, it should be clear that steeper slopes will allow the box to slide downwards faster. Let us try to model and prove this formally.

Suppose that we have another one of these boxes on a separate slope with a steeper incline. We can model this situation by using a smaller width ρ, but using a higher value for the new incline's height σ. We shall similarly enforce $\rho^2 + \sigma^2 = 1$.

For clarity, let the coordinates of the new box be a, b. Following very much the same derivation that we did for the first box, the following system of ODEs can be used to describe the motion of the second box:

$$a' = v_a, b' = v_b$$
$$v_a' = g\rho\sigma, v_b' = -g\sigma^2$$

Suppose that both boxes were initially started at rest and that they start at the same coordinates. We shall prove that the vertical distance traveled by the box on the steeper slope is always greater than that of the other box.

Exercise 10:
To make sure everyone has practice writing down models we shall work through this example together.
Answer: We already know the model of physics: we can just glue the ODEs for both two boxes together.

$$\beta \stackrel{\text{def}}{\equiv} \{x' = v_x, y' = v_y, v_x' = gwh, v_y' = -gh^2, a' = v_a, b' = v_b, v_a' = g\rho\sigma, v_b' = -g\sigma^2, t' = 1\}$$

What should the initial conditions be? We certainly need all of our *constant* assumptions. Always remember to write these down: KeYmaera X and dL formulas/sequents do not know what assumptions you are making on constants unless they are written down. We will also add in our assumption that the new slope is steeper i.e., $\sigma > h$:

$$\Gamma \stackrel{\text{def}}{\equiv} g > 0, h^2 + w^2 = 1, h > 0, w > 0, \sigma^2 + \rho^2 = 1, \sigma > 0, \rho > 0, \sigma > h$$

We will also need some initial assumptions about the positions of the boxes. In order for the competition to be fair, let us just assume that they start at the same coordinates at rest:

$$Init \stackrel{\text{def}}{\equiv} t = 0, x = x_0, y = y_0, v_x = 0, v_y = 0, a = x_0, b = y_0, v_a = 0, v_b = 0$$

Finally, we need to write down a postcondition for this system. Remember that writing down postconditions that clearly correspond to what we want makes your model easier to understand.

$$Safe \stackrel{\text{def}}{\equiv} y \geq b$$

This is what we will want to prove:

$$\Gamma, Init \vdash [\beta]y \geq b$$

In contrast our earlier question for the single slope model, this question is a lot simpler and we will be able to tackle it with straightforward dI,dC. With the foresight of our earlier proof, we can do a rough dI calculation first, which tells us that we need to show $v_y \geq v_b$. This proves easily with dI:

$$\dfrac{\mathbb{R}\ \dfrac{*}{\Gamma \vdash -gh^2 \geq -g\rho^2}}{[':=]\ \dfrac{\Gamma \vdash [v_y' := -gh^2][v_b' := -g\rho^2]v_y' \geq v_b'}{\text{dI}\ \Gamma, Init \vdash [\beta]v_y \geq v_b}}$$

Thus, a dC completes the proof our our desired property:

$$\dfrac{\mathbb{R}\ \dfrac{*}{\Gamma, v_y \geq v_b \vdash v_y \geq v_b}}{[':=]\ \dfrac{\Gamma, v_y \geq v_b \vdash [x_y' := v_y][x_b' := v_b]v_y' \geq v_b'}{\text{dI}\ \dfrac{\Gamma, Init \vdash [\{\beta\,\&\, v_y \geq v_b\}]y \geq b}{\text{dC}\ \Gamma, Init \vdash [\beta]y \geq b}}}$$

Another interesting question is how to we can choose the slopes so that the box moves the greatest distance horizontally. This is not as obvious physically: if the slope were completely flat, then the box would not be moving. On the other hand, if the slope were completely vertical, then the box would drop straight to the floor but not move very far horizontally.

Let us try and see if dI can give us some hints. Suppose that we want to pick inclines that allow us to show $x \geq a$. Using the argument we had above, we would need to first show $v_x \geq v_a$, and thus $gwh \geq g\rho\sigma$.

Recall that $h^2 + w^2 = 1$ and $h > 0$ so we may rewrite h with $h = \sqrt{1-w^2}$ (and similarly for ρ, σ). In other words, we only need to find the maximum value of the function $f(x) = x\sqrt{1-x^2}$ for $0 \leq x \leq 1$. This maximum value is attained at $x = \frac{1}{\sqrt{2}}$, so the maximum horizontal speed is attained with $h = w = \frac{1}{\sqrt{2}}$, i.e., a 45 degree incline.

3.3 One Slippery Slope with a Spring

Note: This section explains a more advanced model for the slippery slope with an addition of a spring.

Let us change our single box model further and suppose that the box is now attached to a spring that acts on the box parallel to the slope. The spring is initially at rest.

Exercise 11:
How can we extend our ODEs to model this situation?
Answer: Reversing the trigonometric calculations we did earlier, we know that if the box has traveled a horizontal distance $x - x_0$ along the slope, then its distance traveled along the slope is given by $\frac{x-x_0}{w}$.

Therefore, we can model the restoring acceleration due to the spring by modifying our differential equations for velocity:

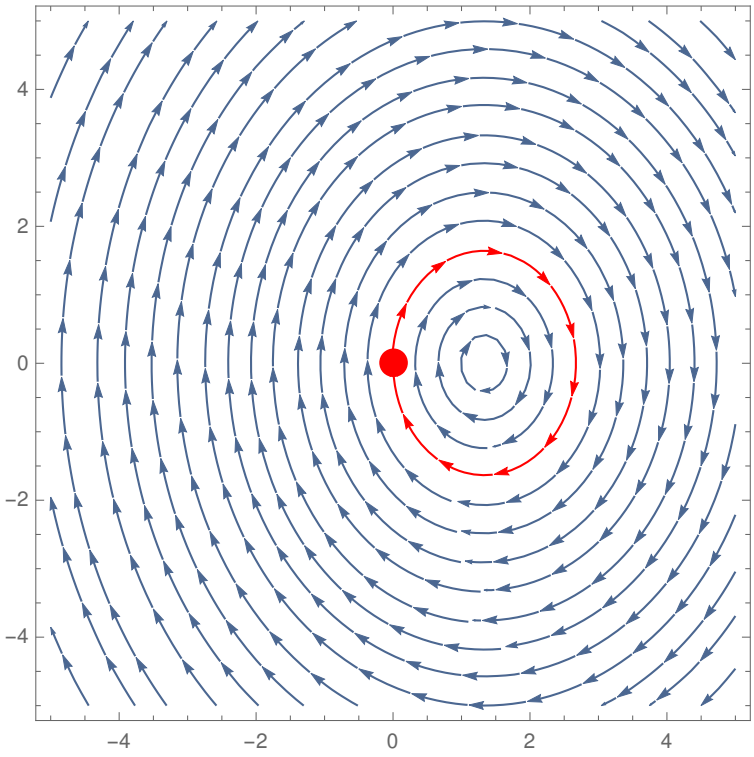
$$v_x' = w\left(gh - k\frac{x-x_0}{w}\right), v_y' = -h\left(gh - k\frac{x-x_0}{w}\right)$$

This quickly becomes a huge mess, so let us focus on studying the x direction only and set $x_0 = 0$. The ODEs describing motion in the horizontal direction can be written as:

$$x' = v_x, v_x' = c - kx$$

where $c = wgh > 0$ is a positive constant, and $k > 0$ is the spring constant.

We may be interested in bounding the horizontal position of the box and perhaps its horizontal velocity. To do this, observe that this simplified system actually describes an oscillator so it will oscillate about the center $v_x = 0, x = \frac{c}{k}$. This is evident once the ODEs are plotted; . Here is the velocity-position plot for $c = 2, k = \frac{3}{2}$. The velocity is plotted vertically while position is plotted horizontally.



From the plot, it is easy to see that the maximum horizontal position of the box is $x \leq 2\frac{c}{k}$. If we tried to prove this right away with dI, we would fail. The technique we have been using so far of cutting in extra invariants would not work either: it tells us to cut in $v \leq 0$, which would not be provable for the above system.

Exercise 12:
How else should we prove this?
Answer: Recall back to Recitation 5 where we actually already encountered a similar situation. We can describe the ellipse by:

$$\frac{(x - \frac{c}{k})^2}{(\frac{c}{k})^2} + \frac{v_x^2}{\frac{c^2}{k}} = 1$$

This implies, in particular that the maximum horizontal position is given by $x = 2\frac{c}{k}$, while the maximum horizontal speed is given by $|v_x| = \sqrt{\frac{c^2}{k}}$.

Exercise 13:
Work through the dI calculation and convince yourself that it works.

Notice that the final approach we discussed at the start of this lecture really shines here. Instead of doing all of the above calculations for the y coordinate again, we can simply use the relationship from Equation 3 to obtain bounds on y.

4 Some Common Questions

In view of the upcoming midterm, we went over some common mistakes and questions.

4.1 Proof Rules vs Axioms

The following note, written by Jonathan Laurent, makes the difference quite clear.

"Let's take this axiom for example:

$$(H \to P) \to [x' = f(x)\&H]P$$

The most reliable way to determine whether or not an axiom is sound is to try and prove it sound. If you succeed, you're done. If you block somewhere, you can try and analyze why and then derive a counterexample from this insight.

So let's try and prove that the axiom above is sound.

1. Let's consider a state s in which $H \to P$ is true.
2. We want to prove that $[x' = f(x)\&H]P$ is also true in state s:
3. Consider a state s' such that (s, s') is a valid transition for $x' = f(x)\&H$.
4. We must prove that P is true in s':
5. By virtue of being in the domain constraint, H is true in s'.
6. Moreover, we have $H \to P$ by assumption.
7. Therefore, P is true in s' and we are done.

Now, take some time to read this proof. Are you happy with it?

Your answer should be no. And the big flaw in this proof is to be found at line 6. Indeed, our assumption is only that $H \to P$ holds in state s. But it might not hold in s'. And so we cannot conclude anything. The differential dynamic logic proof calculus prevents you from drawing incorrect inferences with catastrophic consequences for your CPS. But if you jump to conclusions and incorrectly argue for the soundness of new axioms or rules your reasoning would still be flawed. Logic has the tools to distinguish whether a formula is true in one state, or whether it is valid so true in all states, or whether it's just a formula that isn't true anywhere at all. When arguing for soundness or unsoundness you should use those tools.

This is the major difference between axiom $(H \to P) \to [x' = f(x)\&H]P$, which is unsound, and the similar proof rule:

$$\text{(R12)} \quad \frac{\vdash H \to P}{\vdash [x' = f(x)]P}$$

This rule is actually sound (it can be derived from dW). And the main difference here is that in the premise, $H \to P$ is assumed to be true IN ALL STATES, because a proof rule is sound when the validity of all its premises implies the validity of its conclusion. "

4.2 Finding Counterexamples

A reliable way to generate counterexamples is by using the failed proof to guide you. For a concrete example, we can look at the next part of the note quoted above.

" But let's get back to $(H \to P) \to [x' = f(x)\&H]P$. Now that we suspect that this is unsound, we must find a counterexample. Looking at where the soundness proof failed, we must find H and P such that $H \to P$ is true before a run of the ODE program and false after.

In order to make sure implications do not trip you up, let's first look for a counterexample with H=true. So we must find P along with an ODE program such that P can be true initially and become false after executing the ODE program. This is not hard to find and so you can answer something like:

> Let's take H =true, $P = (x = 0)$ and $f(x) = 1$. Our axiom becomes: $x = 0 \to [x' = 1]x = 0$, which is clearly not valid.

$\cdots$ you should always aim at finding counterexamples that are as simple as possible. "

4.3 Assignment Axiom

In assignment 2, a common mistake was to substitute assignments in the wrong order. This can lead to variable capture (which you will learn about soon) and incorrect results. Until you learn to identify bound variables in chapter 11, it is important to apply the assignment axiom substituting in the assignment closest to the formula first.