Universitat Autònoma de Barcelona

Physics BSc. Undergraduate Thesis

June 2021

Title

Subtitle

Arnau Mas Dorca

Supervised by

Dr Jordi Gaset

Contents

Preface

Preface

Acknowledgements Acknowledgements

Chapter 1

The Herglotz variational problem

We claim that a way of describing non-conservative systems with a variational principle is by what we will call an action-dependent Lagrangian. In all of the standard examples from classical mechanics, the Lagrangian is assumed to be a function of the positions and velocities of the particle. An *action-dependent Lagrangian* is a Lagrangian that is also allowed to depend on the action of the path. However, recall that the action is defined as the integral of the Lagrangian function along the path so this seems awfully circular! Thus, great care has to be taken when formulating the theory of action-dependent Lagrangians, in order to circumvent issues of circularity. The problem of finding the stationary paths of the action functional corresponding to an action-dependent Lagrangian is known as the Herglotz variational problem. In this chapter we will first discuss an *implicit approach* to this problem and some of its shortcomings. Then we will frame the Herglotz problem as a constrained optimisation problem and show how this different approach generalises to field theory.

1.1 The implicit formulation of the Herglotz problem

Let's briefly describe what we will refer to as the *implicit formulation of the Herglotz problem*, as presented in [Laz+18]. We start by writing down a naive equation for the action corresponding to an action dependent Lagrangian. Namely, given a path $q^\mu\colon [a,b] \to M$ we would write something like

$$S[q^\mu] = \int_a^b L(q^\mu(t), \dot{q}^\mu(t), S(t))\,\mathrm{d}t$$

where $S(t)$ is the action of the path until time t. Of course this makes no sense since we are defining S on the left-hand side, and it appears on the right-hand side! However, if, given a path then we define its action as a function of time then we could write something like

$$S[q^\mu](t) = \int_a^t L(q^\mu(s), \dot{q}^\mu(s), S[q^\mu](s))\,\mathrm{d}s \tag{1.1}$$

and if we differentiate with respect to time we actually get an ODE for $S[q^\mu]$! Indeed

$$\dot{S}[q^\mu](t) = L(q^\mu(t), \dot{q}^\mu(t), S[q^\mu](t)). \tag{1.2}$$

Notice that Equation (1.1) actually forces the initial condition $S[q^\mu](a) = 0$. We can even drop this requirement, since all we are interested in is the difference of values of S:

$$S[q^\mu](b) - S[q^\mu](a) = \int_a^b L(q^\mu(t), \dot{q}^\mu(t), S[q^\mu](t))\,\mathrm{d}t. \tag{1.3}$$

What we have here is a functional which, for every path, is defined by an ODE. To find the stationary paths of this functional we would, in principle, have to solve Equation (1.2) for any possible path and among all of them find which ones yield extrema. This is the *Herglotz variational problem*. This approach is in general not feasible. However, just like the variational problem of classical Lagrangian mechanics can be turned into a set of ODEs, the Euler-Lagrange equations, so can the Herglotz problem be turned into a set of ODEs. These are known as the Herglotz equations, which can be written down as

$$\frac{\partial L}{\partial q^\mu} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^\mu} + \frac{\partial L}{\partial \dot{q}^\mu}\frac{\partial L}{\partial z} = 0 \tag{1.4}$$

where the Lagrangian is $L(q^\mu, \dot{q}^\mu, z)$, z being the action dependence. These equations differ from the Euler-Lagrange equations only by one term. And in fact, if L is action-independent, thus $\frac{\partial L}{\partial z} = 0$, we recover exactly the Euler-Lagrange equations.

See [Laz+18] §3.2 of [LL21] or for a derivation of the Herglotz equations following the implicit approach.

1.1.1 Example: the damped harmonic oscillator

Before we move forward, let's see what kind of equations of motion we get from the Herglotz equations. The simplest action-dependence we can introduce —other than no action-dependence at all, of course— is a term linear in z. Let's see what happens when we add this term to the simple harmonic oscillator. Explicitly, consider the Lagrangian

$$L(q, \dot{q}, z) = \tfrac{1}{2}m\dot{q}^2 - \tfrac{1}{2}m\omega^2 q^2 - \gamma z.$$

A straightforward computation shows that for this Lagrangian, Equation (1.4) becomes

$$-m\omega^2 q - m\ddot{q} - \gamma m\dot{q} = 0$$

which after some rearranging can be turned into

$$\ddot{q} + \gamma\dot{q} + \omega^2 q = 0,$$

which one recognises as the equation of motion of the damped harmonic oscillator. The Herglotz principle delivers in its promise: we have just derived the equations of motion of a fundamentally non-conservative system from a variational problem!

1.2 The Herglotz problem as a constrained optimisation problem

We have derived the Herglotz equations, so it would seem we have already solved the theory of action-dependent Lagrangians. However, general relativity is a field theory, so if we wish to understand an action-dependent variant of it we have to know the form the Herglotz equations take for field theory. If we try to apply the implicit method to field theory we run into a number of problems. For one, it is just not very elegant. The action functional is defined implicitly through an ordinary differential equation, one for every path. What's more, this equation is no longer an ODE in field theory, but rather it becomes a PDE. PDEs are notoriously much more difficult to solve thant ODEs, so we would like a way to circumvent this issue.

We can achieve this if we frame the Herglotz problem as a constrained optimisation problem. Not only that, we also get a much clearer and more elegant formulation. We describe what this looks like for mehcanics. First, instead of considering paths in spacetime, $q^\mu\colon [a,b] \to M$, we enlarge the configuration space with one extra quantity, which we will call z. At this point z simply tracks a quantity that changes along the path, but we will later require that z actually match the action of the path at each time. This will be the constraint.

So, we have paths of the form $(q^\mu, z)\colon [a,b] \to M \times \mathbb{R}$. We define a functional on these paths as

$$S[q^\mu, z] = z(a) - z(b) = \int_a^b \dot{z}(t)\,\mathrm{d}t, \tag{1.5}$$

so S is just the change in z along the trajectory. This functional as it stands has no stationary paths, since we can find trajectories with arbitrarily large changes in z, both positive and negative. So we constrain the possible paths. Namely, we require that z

actually represent the action. So we try to find the paths that extremise S only among those that satisfy the constraint

$$\dot{z}(t) = L(q^\mu(t), \dot{q}^\mu(t), z(t)), \tag{1.6}$$

where L is the action-dependent Lagrangian that describes the system. Notice that this is very similar to Equation (1.2).

Say(q^μ, z) is a trajectory that satisfies Equation (1.6). Then

$$S[q^\mu, z] = z(a) - z(b) = \int_a^b \dot{z}(t)\,\mathrm{d}t = \int_a^b L(q^\mu(t), \dot{q}^\mu(t), z(t))\,\mathrm{d}t.$$

So, for paths that satisfy the constraint, the functional S is indeed the action functional, understood as the integral of the Lagrangian along the path.

This is all well and good, but how does one actually go about solving a constrained optimisation problem? It turns out, we can use an infinite dimensional analog of Lagrange multipliers to turn this into a regular optimisation problem to which we can apply the tools of the calculus of variations.

Recall, given some function $f\colon \mathbb{R}^n \to \mathbb{R}$, $x \in \mathbb{R}^n$ is an extremum of f subject to m constraints $g_k\colon \mathbb{R}^n \to \mathbb{R}$ —i.e. $g_k(x) = 0$ — if and only if there exist numbers $\lambda_k \in \mathbb{R}$ such that x is an extremum of the function

$$F = f - \sum_{k=1}^{m} \lambda_k g_k \tag{1.7}$$

without any constraints. The numbers λ_k are called the Lagrange multipliers. It can be shown that this result generalises to infinite dimensional spaces and infinitely many constraints. In our case, the function we are trying to find the extrema of is the functional S. Our constraints are parameterised by $t \in [a, b]$:

$$g_t[q^\mu, z] = \dot{z}(t) - L(q^\mu(t), \dot{q}^\mu(t), z(t)).$$

Thus, replacing sums with integrals in Equation (1.7), the extrema of Equation (1.5) subject to Equation (1.6) will be those that extremise the following functional:

$$\begin{aligned}\tilde{S}[q^\mu, z] &= S[q^\mu, z] - \int_a^b \lambda_t g_t[q^\mu, z]\,\mathrm{d}t \\ &= \int_a^b \dot{z}(t)\,\mathrm{d}t - \int_a^b \lambda_t\Big[\dot{z}(t) - L(q^\mu(t), \dot{q}^\mu(t), z(t))\Big]\,\mathrm{d}t \\ &= \int_a^b (1 - \lambda_t)\dot{z}(t) + \lambda_t\Big(L(q^\mu(t), \dot{q}^\mu(t), z(t))\Big)\,\mathrm{d}t. \end{aligned} \tag{1.8}$$

A couple of observations. First, we are thinking of the Lagrange multipliers as real numbers parameterised by t, but we could equivalently think of them as a function of t and write $\lambda(t)$ instead of λ_t. We will do this. Secondly, and more importantly, if we write the integrand of Equation (1.8) as

$$\tilde{L}(q^\mu(t), \dot{q}^\mu(t), z(t), \dot{z}(t)) = \Big(1 - \lambda(t)\Big)\dot{z}(t) + \lambda(t)\Big(L(q^\mu(t), \dot{q}^\mu(t), z(t))\Big)$$

one should be able to recognise $\tilde{S}$ as something that looks just like a regular old action functional defined by the integral of a regular old Lagrangian, except it is now defined on expanded trajectories (q^μ, z). So we should be able to use the Euler-Lagrange equations to write down the equations of motion of its extremal paths! The equation for z reads

$$0 = \frac{\partial \tilde{L}}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \tilde{L}}{\partial \dot{z}} = \lambda \frac{\partial L}{\partial z} + \dot{\lambda}$$

or equivalently

$$\dot{\lambda} = -\lambda \frac{\partial L}{\partial z}. \tag{1.9}$$

The Euler-Lagrange equations for the positions then are

$$0 = \frac{\partial \tilde{L}}{\partial q^\mu} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \tilde{L}}{\partial \dot{q}^\mu} = \lambda \frac{\partial L}{\partial q^\mu} - \dot{\lambda}\frac{\partial L}{\partial \dot{q}^\mu} - \lambda \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^\mu}$$

and after substituting in Equation (1.9) and dividing through by λ we find

$$0 = \frac{\partial L}{\partial q^\mu} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^\mu} + \frac{\partial L}{\partial z}\frac{\partial L}{\partial \dot{q}^\mu} \tag{1.10}$$

which are exactly the Herglotz equations.

1.3 Action-dependent field theory

We have seen how to derive the Herglotz equations in a more elegant way using Lagrange multipliers. More importantly, we have written down a modified Lagrangian which gives rise to them through standard calculus of variations techniques. This will be very useful as, in some cases (general relativity being one of them), it is easier to derive the equations of motion of a system by direct variation of the action, rather than writing down the Euler-Lagrange equations. We will be able to do exactly this once we have the field theoretic version of Equation (1.8) in our hands.

1.3.1 Classical field theory and Lagrangian densities

First off, we need to set the stage for Lagrangian field theory. The parameter space is no longer just time, but rather all of spacetime, M. Fields are the assignment of some value to each point in spacetime, so we could have scalar fields, vector fields, or, as is the case in general relativity, tensor fields. Let us first fix some notation. We will denote by ϕ some field configuration. If ϕ is a scalar field then it carries no indices. If ϕ is a vector field then it carries one upper index, if it is a tensor field then it carries multiple indices. The metric carries two lower indices since it is a $(0, 2)$ tensor field. In almost all that follows we will assume ϕ is a vector field, but the results we find transfer to tensors of other rank. Of course ϕ depends on spacetime, which we will sometimes write explicitly as $\phi^a(x^\mu)$. As a convention, we will use latin indices for the indices of the field, and reserve greek indices for spacetime coordinates.

The Lagrangian of a field theory is a function of the field and of its derivatives. It is not, however, a function of real values, but rather an n-dimensional —where n is the dimension of spacetime—, differential form, also known as a top form. This is because, the action is defined as the integral of the Lagrangian over some region of spacetime, and the objects we can integrate over spacetime are precisely the top forms.

It turns out, however, that any two top forms differ only by an overall factor, i.e., given two top forms ω_1 and ω_2, there exists a scalar function of spacetime M such that $\omega_1 = f\omega_2$. This means, that if we think of the Lagrangian as some top form $\mathcal{L}$, once we pick a distinguished top form there is a unique scalar function L such that $\mathcal{L} = L\omega$. This distinguished top form will in most cases be the top form induced by the coordinates we are working in, which we will write as $\mathrm{d}^n x$. Sometimes we will use Lagrangian density to refer to $\mathcal{L}$ and Lagrangian to refer to L.

The setup in classical field theory is as follows. Given some Lagrangian, which encodes the system we are studying, we define the action functional on all the possible field configurations as

$$S[\phi^a] = \int_D \mathcal{L}(\phi^a(x^\mu), \partial_\mu \phi^a(x^\mu)) = \int_D L(\phi^a(x^\mu), \partial_\mu \phi^a(x^\mu))\, \mathrm{d}^n x,$$

where D is some region of spacetime where this integral makes sense.

Using the calculus of variations one can show that the stationary configurations of this action functional satisfy the Euler-Lagrange equations of field theory

$$\frac{\partial L}{\partial \phi^a} - \partial_\mu \frac{\partial L}{\partial(\partial_\mu \phi^a)} = 0$$

where a summation convention is implied over μ.

1.3.2 The action flux

How do we generalise this to an action dependent Lagrangian? The most naive approach would be to try to replicate the Herglotz equations from mechanics wholesale and write down

$$\frac{\partial L}{\partial \phi^a} - \partial_\mu \frac{\partial L}{\partial(\partial_\mu \phi^a)} + \frac{\partial L}{\partial(\partial_\mu \phi^a)} \frac{\partial L}{\partial z} = 0.$$

However this will not work because the last term has a pesky free μ index. This seems to suggest that we need to modify the nature of z. We had claimed before that z represented the action along the path, but if we look at Equation (1.5) we see this is not quite right. In mechanics the analog of D is $[a,b]$. The difference $z(a) - z(b)$ can also be written as $\int_{\partial[a,b]} z$, since the boundary of $[a,b]$ is just a and b. This seems to indicate that the correct analog of Equation (1.5) for field theory should be

$$S[\phi^a, z] = \int_{\partial D} z.$$

What kind of obejct should z be then? ∂D has dimension $n-1$, so z has to be something we can integrate over an $(n-1)$-dimensional manifold, i.e. a differential $(n-1)$-form. As it turns out, $(n-1)$-forms behave almost like vector fields. Specifically, by expanding in the basis $\mathrm{d}x_\mu$ we have $z = z^\mu \, \mathrm{d}x_\mu$. Thus, what z represents is the action flux. We encode this in a constraint perfectly analogous to Equation (1.6),

$$\mathrm{d}z(x^\mu) = \mathcal{L}(\phi^a(x^\mu), \partial_\mu \phi^a(x^\mu), z^\mu(x^\mu)). \tag{1.11}$$

$\mathrm{d}z$ is the exterior derivative of z, which is a top form, so this equality makes sense. In coordinates it is easy to show that $\mathrm{d}z = \partial_\nu z^\nu \, \mathrm{d}^n x$, so that Equation (1.11) reads in coordinates as

$$\partial_\nu z^\nu(x^\mu) = L(\phi^a(x^\mu), \partial_\mu \phi^a(x^\mu)). \tag{1.12}$$

This expresses the fact that $\mathrm{d}z$ has to be the action density. Indeed, for field configurations that satisfy this constraint then one has, applying Stokes' theorem

$$S[\phi^a, z^\mu] = \int_{\partial D} z = \int_D \mathrm{d}z = \int_D \mathcal{L}(\phi^a, \partial_\mu \phi^a, z^\mu).$$

So we have arrived at the right formulation of the Herglotz variational problem for field theory.

1.3.3 Constrained optimisation in field theory

So just like before, we will turn this constrained optimisation problem into an unconstrained one using Lagrange multipliers. The expanded action, in analogy with Equation (1.8) will be

$$\tilde{S}[\phi^a, z^\mu] = \int_D (1-\lambda)\,\mathrm{d}z + \lambda\mathcal{L}(\phi^a, \partial_\mu\phi^a, z^\nu) = \int_D \mathrm{d}^n x\Big[(1-\lambda)\partial_\mu z^\mu + \lambda L(\phi^a, \partial_\mu\phi^a, z^\nu)\Big] \tag{1.13}$$

Note that the Lagrange multiplier λ is a function of spacetime. Let us write down the integrand of Equation (1.13) as an expanded Lagrangian:

$$\tilde{\mathcal{L}}(\phi^a, \partial_\mu\phi^a, z^\nu, \partial_\mu z^\nu) = \tilde{L}(\phi^a, \partial_\mu\phi^a, z^\nu, \partial_\mu z^\nu)\,\mathrm{d}^n x = \Big[(1-\lambda)\partial_\mu z^\mu + \lambda L(\phi^a, \partial_\mu\phi^a, z^\nu)\Big]\,\mathrm{d}^n x. \tag{1.14}$$

1.3.4 The Herglotz equations for field theory

Finally, we can derive the Herglotz equations for field theory from the expanded Lagrangian we have just obtained in Equation (1.14). The equations for the action flux are

$$0 = \frac{\partial\tilde{L}}{\partial z^\nu} - \partial_\mu\frac{\partial\tilde{L}}{\partial(\partial_\mu z^\nu)} = \lambda\frac{\partial L}{\partial z^\nu} + \partial_\mu(\lambda\delta^\mu_\nu) = \lambda\frac{\partial L}{\partial z^\nu} + \partial_\nu\lambda.$$

So, rearranging,

$$\partial_\nu\lambda = -\lambda\frac{\partial L}{\partial z^\nu}. \tag{1.15}$$

And for the values of the field

$$0 = \frac{\partial\tilde{L}}{\partial\phi^a} - \partial_\mu\frac{\partial\tilde{L}}{\partial(\partial_\mu\phi^a)} = \lambda\frac{\partial L}{\partial\phi^a} - (\partial_\mu\lambda)\frac{\partial L}{\partial(\partial_\mu\phi^a)} - \lambda\partial_\mu\frac{\partial L}{\partial(\partial_\mu\phi^a)}$$

and, when plugging in Equation (1.15) and dividing through by λ we arrive at the field theoretical Herglotz equations

$$\frac{\partial L}{\partial\phi^a} - \partial_\mu\frac{\partial L}{\partial(\partial_\mu\phi^a)} + \frac{\partial L}{\partial z^\mu}\frac{\partial L}{\partial(\partial_\mu\phi^a)} = 0. \tag{1.16}$$

Chapter 2

Bibliography

[Laz+18] Matheus J. Lazo et al. "An action principle for action-dependent Lagrangians: Toward an action principle to non-conservative systems". In: *Journal of Mathematical Physics* 59.3 (Mar. 2018), p. 032902. ISSN: 0022-2488. DOI: 10.1063/1.5019936. URL: https://aip.scitation.org/doi/10.1063/1.5019936 (visited on 03/30/2021).

[LL21] Manuel de León and Manuel Lainz. "A review on contact Hamiltonian and Lagrangian systems". In: *arXiv:2011.05579 [math-ph]* (Feb. 2021). arXiv: 2011.05579. URL: http://arxiv.org/abs/2011.05579 (visited on 03/08/2021).