Lagrangians and complex derivatives

Chris Foster

A common procedure in physics is to make use of complex derivatives of nonanalytic functions. However, the natural definition of complex derivative in terms of limits no longer makes sense in this case. One upshot is that the usual chain rule for complex derivatives is no longer valid, and we obtain something looking identicle to the multivariate chain rule on $\mathbb{R}^2$.

In these notes I explore the issue a bit, starting from a somewhat common culprit in classical field theory.

Warning: Some parts of these notes I now know to be incomplete and perhaps even misleading. I still think there's something useful here however and intend to update them when I have time. See section I for details.

I. WARNING ON INCOMPLETENESS OF THESE NOTES

Since writing these notes I've become aware of ways to make the theory much more capable. In particular, the chain rule can be made to work by introducing an extra operator,

$$\frac{\partial}{\partial z^*} := \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}) \tag{1}$$

In addition to the $\frac{d}{dz}$ (more properly $\frac{\partial}{\partial z}$) defined below.

This *Wirtinger calculus* [1] is described briefly in the excellent book by Remmert [2]. Further useful notes may also be found online at [3].

I plan to update these notes at some stage, but even with the caveats above I hope they can still be a little useful.

II. THE PROBLEM

It is sometimes useful to consider classical field theories of complex rather than real-valued fields. To obtain the equations of motion, physics students are then told to "pretend that the field and its complex conjugate are independent" and to take derivatives with respect to the complex fields. It is not immediately obvious exactly why this procedure works, how it is well-defined, or if it works in any sort of generality at all. This is the original problem which motivated these notes.

Consider a Lagrangian functional, $\mathcal{L}[\psi]$ say, where $\psi\colon \mathbb{R}^3 \to \mathbb{C}$ is a complex-valued field. If any example is going to be familiar, the equation

$$\mathcal{L}[\psi] = \int d^3r\, \frac{i\hbar}{2}\left(\psi^*(r)\dot{\psi}(r) - \dot{\psi}^*(r)\psi(r)\right) - \left(\frac{\hbar^2}{2m}\left|\nabla\phi(r)\right|^2 + V(r)\left|\phi(r)\right|^2\right) \tag{2}$$

will, as it gives rise to the Schrödinger equation as the equation of motion,

$$i\hbar\frac{d\psi}{dt} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi. \tag{3}$$

In general, we get the equations of motion by employing the usual process from Lagrangian mechanics: take derivatives of $\mathcal{L}$ with respect to the independent variables — in this case the field ψ — and put them into the Euler-Lagrange equations. We usually write something like

$$\frac{d}{dt}\left(\frac{\delta\mathcal{L}}{\delta\dot{\psi}(r)}\right) = \frac{\delta\mathcal{L}}{\delta\psi(r)}. \tag{4}$$

where $\frac{\delta\mathcal{L}}{\delta\psi(r)}$ is a *functional* derivative with respect to the field variable 'at the point r'. One way to think about these are as an infinite set of coupled equations parametrised by the continuous variable r.

The confusing part is knowing how to evaluate these derivatives, and exactly what we mean by them. It is often suggested — see for example [4, page 598] — that we should consider ψ and ψ^* to be "independent" for the purposes of taking the derivatives. This prescription is rather confusing, and it's certainly not obvious it works in general — or even why it works at all. I think that there's two main reasons for the confusion:

- Obviously, ψ and ψ^* are *not* independent in the usual sense: change one, and you specify exactly what the other is.
- Real valued functions of complex variables — like the Lagrangian — are not even differentiable in complex analysis (unless they are constant). This is a straightforward application of the Cauchy-Riemann conditions [1]

III. A DERIVATIVE OPERATOR FOR NONANALYTIC FUNCTIONS

One complicating factor above is that the derivatives are functional derivatives with respect to the field ψ. However, the central problems remain if we forget about the fields, and consider a Lagrangian function of a single complex variable, $\mathcal{L}\colon \mathbb{C} \to \mathbb{R}$.

In this case we don't have to worry about the functional differentiation, and can use the usual straightforward definition for complex differentiation (see, for example [5, page 30]):

$$\frac{df}{dz}(z_0) = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}. \tag{5}$$

Recall that we say a function is *analytic at* z_0 if it is differentiable at all points in a neighborhood of z_0, and that it is analytic on some domain D if it is analytic at all points in D.

One consequence of the definition of derivative is that analytic functions f obey the equation

$$\frac{df}{dz} = \frac{\partial f}{\partial x} = -i\frac{\partial f}{\partial y} \tag{6}$$

for $z = x + iy$. As a consequence, we might also choose to write

$$\frac{df}{dz} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right). \tag{7}$$

---

[1] Getting ahead of myself, let me remind people that the Cauchy-Riemann conditions are necessary (but not sufficient) conditions for a complex function, $f\colon \mathbb{C} \to \mathbb{C}$ to have a complex derivative. If $f(z) = u(z) + iv(z)$ with $u, v\colon \mathbb{C} \to \mathbb{R}$, the Cauchy-Riemann conditions are $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, and $\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$.

Although functions which are not analytic don't have a well-defined value for the limit given in equation (5), we may decide to *define* a "derivative" by equation (7). This definition gives us a derivative for any complex function $f(z) = u(z) + iv(z)$ when the real and imaginary components $u = \mathrm{Re}(f)$ and $v = \mathrm{Im}(f)$ are differentiable. Why might this be a convenient definition?

The answer to this is twofold:

- The definition reduces to the correct form in the case that our function is analytic. In this case all the usual rules of complex calculus apply — linearity, the product rule, chain rule etc. However, there are an infinity of such possible definitions; the next property is the real reason for our choice.

- $f(z) = z^*$ is "independent" of z in the sense that

$$\frac{dz^*}{dz} = 0. \tag{8}$$

  Provided the product rule is valid, this allows us to write things like

$$\frac{d}{dz} z^2 z^* = 2zz^*. \tag{9}$$

There is a major catch in using our rather ad-hoc definition for derivative; it doesn't obey some of the rules for calculus which we are so familiar with. In particular the chain rule — and hence the quotient rule — are invalid. On the other hand, our definition is still linear and remains compatible with the product rule, so it's not completely worthless.

To be convinced that the chain rule doesn't work for our derivative, consider the functions $f(z) = z + z^* = 2x$ and $g(z) = z - z^* = 2iy$. Clearly, $f(g(z)) = 0$, so evaluating the derivative directly, we have

$$\frac{df(g(z))}{dz} = \frac{d}{dz} 0 = 0. \tag{10}$$

On the other hand, linearity gives us $f'(z) = 1$ and $g'(z) = 1$, so using the chain rule, we would expect

$$\frac{df(g(z))}{dz} = \left.\frac{df}{dz}\right|_{g(z)} \frac{dg}{dz} = 1. \qquad \text{(wrong!)} \tag{11}$$

The quotient rule is most easily derived using a combination of the chain rule and the product rule, so it's not surprising that it fails as well. This can be seen using the same two functions f and g defined above.

The upshot of all this is that taking these kinds of "derivatives" of nonanalytic functions with respect to complex variables is rather dangerous unless we remember that they don't follow all of the usual rules.

IV. USING THE "DERIVATIVE" TO OBTAIN EQUATIONS OF MOTION

The derivative defined by equation (7) can be used to derive equations of motion from a Lagrangian in a manner which is — formally at least — exactly the same as the usual method. This is provided of course that we can actually calculate the appropriate derivatives without using the chain rule.

To see how it works, consider a Lagrangian of a single complex variable $\mathcal{L}(z)$ as in the previous section. We have

$$\frac{d}{dt}\left(\frac{d\mathcal{L}(z)}{d\dot{z}}\right) = \frac{d\mathcal{L}(z)}{dz} \tag{12}$$

$$\implies \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}} - i\frac{\partial\mathcal{L}}{\partial\dot{y}}\right) = \frac{\partial\mathcal{L}}{\partial x} - i\frac{\partial\mathcal{L}}{\partial y}, \tag{13}$$

and since the Lagrangian is real, we may equate real and imaginary parts. This gives the pair of equations,

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}}\right) = \frac{\partial\mathcal{L}}{\partial x} \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{y}}\right) = \frac{\partial\mathcal{L}}{\partial y}, \tag{14}$$

which are exactly what we would have got if we had treated $\mathcal{L}$ as a real function of two real variables x and y.

V. DISCUSSION

I imagine that the general ideas expressed above carry over to the case of Lagrangian *functionals* rather than functions. Since this isn't really the reason for this note, I degenerate into being rather vague about this point, in the tradition of good physicists.

The take-home message is that the "pretend ψ^* is independent of ψ and differentiate" recipe works well in the situations where the Lagrangian may be differentiated using only linearity and the product rule. In my very limited experience using field theory Lagrangians, this is often the case. In other situations beware — the carefully correct solution is to explicitly write $\mathcal{L}$ in terms of real component fields and differentiate only with respect to these components.

[1] W. Wirtinger. Zur formalen theorie der funktionen von mehr komplexen vernderlichen (on the formal theory of functions of many complex variables). *Mathematische Annalen*, 97(1), December 1927.

[2] Reinhold Remmert. *Theory of Complex Functions*. Springer-Verlag, 1991.

[3] K. Kreutz-Delgado. The complex gradient operator and the CR-calculus. `http://dsp.ucsd.edu/~kreutz/PEI-05%20Support%20Files/complex%_derivatives.pdf`, (last accessed 27/3/08), 2006.

[4] Herbet Goldstein. *Classical Mechanics*. Addison-Wesley, second edition, 1980.

[5] R. Churchill. *Complex Variables and Applications*. McGraw-Hill, second edition, 1960.