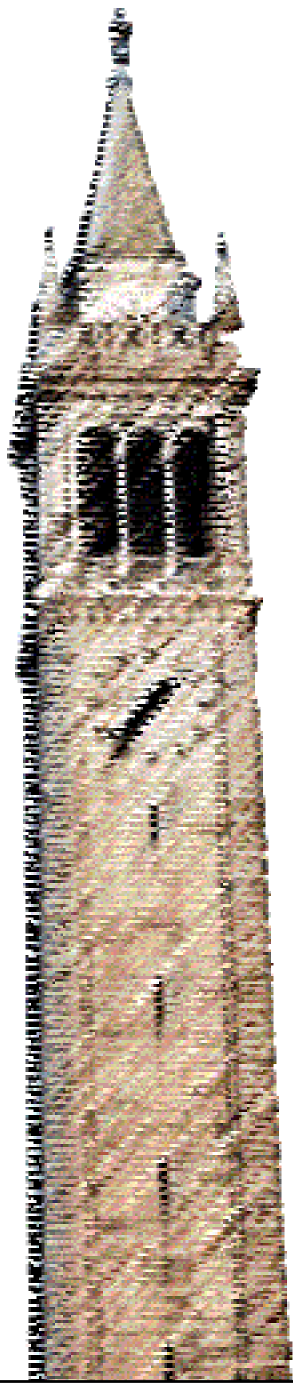
Stationary points of a real-valued function of a complex variable



David G. Messerschmitt

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2006-93

http://www.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-93.html

June 27, 2006

Stationary points of a real-valued function of a complex variable

by
David G. Messerschmitt
University of California at Berkeley

The problem

Let $f(z)$ be a *real-valued* function of a complex variable z. In the course of finding extreme points of $f(z)$, the values of z for which f is maximum or minimum, a useful tool is differential calculus. Since $\left.\frac{\partial f}{\partial z}\right|_{z=z_0}$ is a measure of the slope of f in the neighborhood of $z = z_0$, if f is differentiable then $\left.\frac{\partial f}{\partial z}\right|_{z=z_0} = 0$ is a necessary condition for a maximum or minimum of f at $z = z_0$. Knowing these *stationary points*, we can determine which are a maximum, minimum, or saddle point.

Unfortunately, it is usually the case that this derivative doesn't exist (f is not analytic) because $f(z)$ is implicitly a function of both z and z^*. This challenge arises because differentiability in the complex domain requires that the slope of f is identical for every trajectory in the complex plane through $z = z_0$.

> **Example**: Let $f(z) = |z|^2$. As f is not analytic, we cannot *directly* use the differential calculus to find that the only stationary point is a $z = 0$, even though it is obvious that f is convex and has a single global minimum at $z = 0$.
>
> We can establish that $f(z) = |z|^2$ is not analytic be looking at the limit of
> $$\frac{f(z+\Delta z) - f(z)}{\Delta z} = z \cdot \frac{(\Delta z)^*}{\Delta z} + z^* + |\Delta z|^2 .$$
> The second term is a constant, and the third term goes to zero, so neither is a problem. The problem term is the first (except when it goes away, at $z = 0$), because $(\Delta z)^* / \Delta z$ has unit magnitude (so it doesn't go to zero) but its angle depends on the angle of Δz, and thus the limit of this term as $\Delta z \to 0$ depends on

that angle. To be analytic, the limit of the first term must be independent of the direction (angle) of $\Delta z \to 0$, and that is the case only at $z = 0$.

First approach: formulate in terms of real variables

An approach to avoid this complex differentiability problem is to formulate the problem in terms of two real variables, the real and imaginary parts of z. Let $z = x + i \cdot y$ and (with a minor abuse of notation) $f(z) = f(x, y)$. Now $z = z_0$ is a stationary point when

$$\left.\frac{\partial f}{\partial x}\right|_{z=z_0} = \left.\frac{\partial f}{\partial y}\right|_{z=z_0} = 0 .$$

Example: When $f(z) = |z|^2$, $f(x, y) = x^2 + y^2$ and the conditions for a stationary point become

$$\frac{\partial f}{\partial x} = 2 \cdot x = 0 \text{ and } \frac{\partial f}{\partial y} = 2 \cdot y = 0 .$$

Thus $2x_0 = 2y_0 = 0$, or $z_0 = 0$, is a stationary point. Note that this equivalent real-valued function of two real variables is differentiable, so there are no obstacles to using the calculus to find the stationary point and thus the minimum.

While this method works, it is cumbersome because it involves the extra step of substituting $x + i \cdot y$ for z. Thus, we seek an equivalent method that works directly with the complex variable z.

Second approach: equivalent conditions in terms of complex variables

Assume that f can be represented as

$$f(z) \equiv g(z, z^*)$$

where $g(q, s)$ is an analytic function of *two* complex variables q and s. Because of the assumption of analyticity, we can freely differentiate g w.r.t. either of its arguments. A convenient (and standard) notation for the two partial derivatives is

$$g_q(\cdot,\cdot) \equiv \frac{\partial g}{\partial q} \text{ and } g_s(\cdot,\cdot) \equiv \frac{\partial g}{\partial s} .$$

Example: For $f(z) = |z|^2$, the definition $g(q, s) = q \cdot s$ leads to $g(z, z^*) = f(z)$ and further $g_q(q, s) = s$ and $g_s(q, s) = q$.

More generally, we can simply substitute q for all instances of z and s for all instances of z^* to arrive at the desired representation.

Now, substituting $g(z,z^*)$ for $f(z)$, we can apply the chain rule of differential calculus to find a different representation of the stationary point,

$$\frac{\partial f}{\partial x} = g_q(z,z^*)\cdot\frac{\partial z}{\partial x} + g_s(z,z^*)\cdot\frac{\partial(z^*)}{\partial x} = 0$$
$$\frac{\partial f}{\partial y} = g_q(z,z^*)\cdot\frac{\partial z}{\partial y} + g_s(z,z^*)\cdot\frac{\partial(z^*)}{\partial y} = 0$$

Evaluating the four derivatives, this becomes

$$\frac{\partial f}{\partial x} = g_q(z,z^*) + g_s(z,z^*) = 0$$
$$\frac{\partial f}{\partial y} = i\cdot g_q(z,z^*) - i\cdot g_s(z,z^*) = 0$$

which has a unique solution

$$g_q(z,z^*) = g_s(z,z^*) = 0\,.$$

Convenient shorthand for this (as long as we agree on its precise meaning) is

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial z^*} = 0\,.$$

We have uncovered that the stationary point can be found by taking the partial derivatives of $f(z)$ w.r.t. *both* z and z^*, considering them as *independent* variables, and setting those derivatives to zero.

> **Example**: For $f(z) = |z|^2 = z\cdot z^*$, considered as functions of independent variables z and z^*, the stationary point is $\frac{\partial f}{\partial z} = z^* = 0$ and $\frac{\partial f}{\partial z^*} = z = 0$, or $z = 0$.

Actually, the method we developed works fine for the more general case of a complex-valued function of a complex variable that is a function of both z and z^*, because we have not yet taken account of the more restrictive condition that f is real valued. That is why, in the last example, the two conditions yield redundant information.

> **Example**: An example of a complex-valued function is $f(z) = z^2 + z^*$. The two conditions for a stationary point are
> $$\frac{\partial f}{\partial z} = 2\cdot z = 0 \text{ and } \frac{\partial f}{\partial z^*} = 1 = 0\,.$$
> There is no stationary point for this function because the second condition can never be satisfied. Note, however, that as expected the two conditions are *not* redundant.

> **Example**: Another complex-valued function, this one with a stationary point, is $f(z) = (z-a)(z^*-b)$ for complex constants a and b. Here we get

$$\frac{\partial f}{\partial z} = z^* - b = 0 \text{ and } \frac{\partial f}{\partial z^*} = z - a = 0 .$$

Thus there is a stationary point if $b = a^*$, in which case it falls at $z = a$.

Back to the real-valued case

The more restrictive condition that f is real-valued allows us to reduce the two conditions for a stationary point down to a single condition.

In this case, both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ must be real-valued, since they are derivatives of a real-valued function w.r.t. real variables. Thus, it must be that

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x}\right)^* \text{ and } \frac{\partial f}{\partial y} = \left(\frac{\partial f}{\partial y}\right)^*$$

Plugging in the earlier expressions for these derivatives,

$$g_q(z, z^*) + g_s(z, z^*) = g_q^*(z, z^*) + g_s^*(z, z^*)$$

$$i \cdot g_q(z, z^*) - i \cdot g_s(z, z^*) = -i \cdot g_q^*(z, z^*) + i \cdot g_s^*(z, z^*) ,$$

we solve these equations to arrive at the conclusion that

$$g_q(z, z^*) = g_s^*(z, z^*) .$$

This establishes that the redundancy observed in an earlier example will *always* occur for a real-valued function. Thus, the single condition

$$g_s(z, z^*) = 0$$

suffices to establish a stationary point. In our shorthand, this becomes

$$\frac{\partial f}{\partial z^*} = 0 .$$

Example: Let $f(z) = \alpha \cdot \left|z - a\right|^2 - \beta \cdot |z - b|^2$ for complex constants a and b and real constants α and β. The first step is to write this in terms of z and z^*,

$$f(z) = \alpha \cdot (z - a) \cdot (z^* - a^*) - \beta \cdot (z - b) \cdot (z^* - b^*) .$$

Taking the single derivative w.r.t. z^*, considering z as an independent variable that is held constant,

$$\frac{\partial f}{\partial z^*} = \alpha \cdot (z - a) - \beta \cdot (z - b) = 0 .$$

So long as $\alpha \neq \beta$ there is a stationary point at $z = \frac{\alpha \cdot a - \beta \cdot b}{\alpha - \beta}$. In general this stationary point is complex-valued, as a and b can be complex-valued.

Example: Let $f(z) = |z - a| = \sqrt{(z - a) \cdot (z^* - a^*)}$. Then

$$\frac{\partial f}{\partial z^*} = \frac{(z-a)}{2\sqrt{(z-a)\cdot(z^*-a^*)}} = \frac{(z-a)}{2\cdot|z-a|}$$

Since the magnitude of $(z-a)/|z-a|$ is always unity for $z \neq a$, there are no stationary points for $|z| > a$. Further, the derivative does not exist at $z = a$, so this is not a stationary point either. This is not surprising, since this function is a cone, and is thus "pointed" (its slope is discontinuous) at its minimum $z = a$.

Vector case

These results are easily extended to the case of a function of a vector of complex variables. Let $\vec{Z} = [z_1, z_2, .., z_n]^T$ and assume that $g(\vec{Z}, \overrightarrow{Z^*})$ is an analytic function of complex vectors $\vec{Z}$ as well as its conjugate. Then the conditions for a stationary point are $\nabla_{\vec{Z}}\, g = \vec{0}$ and $\nabla_{\overrightarrow{Z^*}}\, g = \vec{0}$, where ∇ is the gradient operator. For a real-valued function, these conditions are redundant and can be reduced to $\nabla_{\overrightarrow{Z^*}}\, g = \vec{0}$.

Example (observation of a signal in additive noise): Suppose we have L observations

$$O_i = S + N_i,\ 1 \le i \le L$$

where S and N_i, $1 \le i \le L$ are complex-valued zero-mean mutually uncorrelated random variables with variance $\sigma_S{}^2$ and $\sigma_N{}^2$ respectively. Also define a signal-to-noise ratio,

$$SNR = \frac{\sigma_S{}^2}{\sigma_N{}^2}.$$

The objective is to form an estimate of signal S by forming a linear combination of the observations,

$$\hat{S} = \sum_{i=1}^{L} \alpha_i \cdot O_i$$

with the goal of minimizing the mean-square error $E\left|\hat{S} - S\right|^2$,

$$E\left|\hat{S} - S\right|^2 = \left|\sum_{i=1}^{L} \alpha_i - 1\right|^2 \cdot \sigma_S{}^2 + \sum_{i=1}^{N} \left|\alpha_i\right|^2 \cdot \sigma_N{}^2.$$

We would like to choose the α_i to minimize this. As a starting point it is better to write this as

$$E\left|\hat{S} - S\right|^2 = \left(\sum_{i=1}^{L} \alpha_i - 1\right)\cdot\left(\sum_{i=1}^{L} \alpha_i{}^* - 1\right)\cdot \sigma_S{}^2 + \sum_{i=1}^{N} \alpha_i \cdot \alpha_i{}^* \cdot \sigma_N{}^2$$

and setting the partial derivative w.r.t. $\alpha_n{}^*$ equal to zero,

$$\left(\sum_{i=1}^{L} \alpha_i - 1\right)\cdot \sigma_S{}^2 + \alpha_n \cdot \sigma_N{}^2 = 0,\ 1 \le n \le L.$$

From this we conclude that a stationary point occurs where all the coefficients are equal,

$$\alpha_n = \alpha = \left(1 - \sum_{i=1}^{L} \alpha_i\right) \cdot SNR = (1 - L \cdot \alpha) \cdot SNR \text{ , } 1 \le n \le L \text{ ,}$$

or

$$\alpha = \frac{1}{L + \frac{1}{SNR}} .$$

$$E\left|\hat{S} - S\right|^2 = \frac{{\sigma_N}^2}{L + \frac{1}{SNR}} .$$

Note especially the dependence on L: as L increases, adding more observations reduces the variance of the estimation error, especially for $L >> \frac{1}{SNR}$.