APPENDIX A

Review

A.1. Complex numbers

This section is a summary of practical algebra in the complex number system.

Technically, a complex number is an ordered pair of real numbers; in other words, the set $\mathbb{C}$ of complex numbers is the same as R^2, the usual XY plane. Instead of writing a complex number z as (x, y) we will generally write it as $x + iy$. The *real part* of z is x and the *imaginary part* of z is y.

Caution: The imaginary part of a complex number is a **real** number.

Complex numbers of the form $x + 0i = x$ are just real numbers, while complex numbers of the form $0 + iy = iy$, with $y \neq 0$, are called *pure imaginary* numbers.

In this form the rules of algebra, for addition, subtraction, and multiplication can be summarized as:

(1) Use the commutative, associative and distributive laws as for real numbers.
(2) Remember that 0 and 1 are units for addition and multiplication, respectively.
(3) Use $i^2 = -1$ to reduce powers of i.

For example, $i^3 = i^2 \cdot i = -1 \cdot i = -i$. Also,

$$\begin{aligned}(2+i)(3-2i) &= 2\cdot 3 + 2\cdot(-2i) + i\cdot 3 + i\cdot(-2i)\\ &= 6 - 4i + 3i - 2i^2 = 6 - i - 2\cdot(-1) = 8 - i.\end{aligned}$$

There is an important operation for complex numbers called *conjugation*, which is defined by "changing the sign of the imaginary part". Specifically, if $z = x + iy$ then $\overline{z} = x - iy$. The main properties of congugation are:

(1) $\overline{(z \pm w)} = \overline{z} \pm \overline{w}$.
(2) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.
(3) $\overline{\overline{z}} = z$.
(4) z is real if and only if $z = \overline{z}$.
(5) The real and imaginary parts of z are $\dfrac{1}{2}(z + \overline{z})$ and $\dfrac{1}{2i}(z - \overline{z})$.
(6) $z \cdot \overline{z} = x^2 + y^2$. This is real and non-negative, and is 0 if and only if $z = 0$.

Using the conjugate we can calculate multiplicative inverses, so we can divide. The idea is to "rationalize the denominator" by multiplying top and bottom by the conjugate of the denominator. For example,

$$\frac{2+i}{3+2i} = \frac{(2+i)(3-2i)}{(3+2i)(3-2i)} = \frac{8-i}{3^2-(2i)^2} = \frac{8-i}{9-(-4)} = \frac{8}{13} - \frac{i}{13}.$$

We can also use conjugation to define the *absolute value* or *modulus* of a complex number as $|z| = \sqrt{z \cdot \overline{z}} = \sqrt{x^2+y^2}$. For example, $|3+2i| = \sqrt{9+4} = \sqrt{13}$.

Caution: The formula for $|z|$ is $\sqrt{x^2+y^2}$, **not** $\sqrt{x^2+(iy)^2}$. Remember that y is real.

The absolute value has many of the same properties as the absolute value of real numbers:

(1) $|zw| = |z| \cdot |w|$.
(2) $|z| \geq 0$, and $|z| = 0$ if and only if $z = 0$.
(3) $|z+w| \leq |z| + |w|$.

Here are a few things to be careful about:

(1) $|z|^2 = z^2$ is true for real numbers, but is **not** generally true for complex numbers.
(2) $|z| = \pm z$ is true for real numbers, but is **not** generally true for complex numbers.

There is no useful way to define an order relation for complex numbers; so $z \leq w$ has no meaning. All we can do is compare magnitudes, using absolute values.

Instead of using Cartesian coordinates, we can represent complex numbers using polar coordinates. The usual formulas relating Cartesian and polar coordinates are

$$x = r\cos\theta \qquad r = \sqrt{x^2+y^2}$$
$$y = r\sin\theta \qquad \theta = \arctan\left(\frac{y}{x}\right)$$

Applying them to $z = x + iy$, we see that $z = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta)$, where $r = |z|$. The polar angle θ is called the *argument* of z. It can be any angle if $z = 0$, and otherwise it is defined up to the addition of an integer multiple of 2π. If we multiply two complex numbers in polar coordinates and remember the addition formulas for the sine and cosine, we are led to

$$r_1(\cos\theta_1 + i\sin\theta_1) \cdot r_2(\cos\theta_2 + i\sin\theta_2) = r_1 r_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)).$$

That is, when you multiply two complex numbers you multiply their absolute values, but you add their arguments. In geometric terms this means that multiplication of points in the plane (i.e., complex numbers) by a fixed complex number z is the transformation defined by a scale change of $r = |z|$, followed by a rotation around the origin through the angle θ.

We can apply this to calculating z^n. We write $z = r\cos\theta + i\sin\theta$ in polar coordinates. Then $|z^n| = |z|^n = r^n$, and the argument of z^n is obtained by adding n

copies of the argument of z, to give $n\theta$. So we have

$$z^n = (r(\cos\theta + i\sin\theta)) = r^n\left(\cos(n\theta) + i\sin(n\theta)\right) \tag{A.1}$$

This is known as *De Moivre's formula.*

Complex numbers were introduced to solve polynomial equations – indeed, the quadratic equation often gives non-real solutions for quadratic equations. There are similar, but much more complicated, formulas for solving third and fourth degree polynomial equations, and these also involve complex numbers. Mathematicians suspected for many years that complex numbers were enough to solve all polynomial equations. This was first proved by Gauss in 1799, although there were some gaps in his proof; several different completely rigorous proofs have since appeared. It is important enough to be called the Fundamental Theorem of Algebra:

THEOREM A.1 (Fundamental Theorem of Algebra). *If $P(z)$ is a polynomial of degree $m > 0$ with complex coefficients then $P(z)$ has exactly m ccomplex roots $\alpha_1, \alpha_2, \dots, \alpha_m$ (listed with multiplicity). In other words, $P(z)$ factors as $c(z-\alpha_1)(z-\alpha_2)\dots(x-\alpha_m)$ where c is the coefficient of z^m in $P(z)$.*

Even if the original polynomial has real coefficients there will probably be non-real roots. The following gives a way of organizing them.

PROPOSITION A.2. *If $P(z)$ is a polynomial of degree $m > 0$ with real roots then the roots of $P(z)$ can be listed as $r_1, r_2, \dots, r_p, c_1, \overline{c}_1, c_2, \overline{c}_2, \dots c_q, \overline{c}_q$ (listed with multiplicity) where $p + 2q = m$, the roots r_i are real, and the roots c_i and $\overline{c}_i$ are non-real. In other words, complex roots occur in conjugate pairs.*

PROOF. The key observation is that $\overline{P(z)} = P(\overline{z})$, which is true since conjugation preserves sums and products and leaves the coefficients of $P(z)$ unchanged, since these coefficients are real. From this we see that if $P(z) = 0$ then $P(\overline{z}) = 0$. It is then clear that each non-real root z is matched up with its conjugate $\overline{z}$, and so we can list the roots as in the proposition. □

A.2. Partial derivatives

For a function f of two or more variables the *partial derivatives* are defined by differentiating with respect to one variable, while treating all other variables as constants. For example, if $f(x, y) = x^2 + y^3 - \sin(3x+5y)$ then the partial derivatives are

$$\frac{\partial f}{\partial x} = 2x - 3\cos(3x + 5y), \quad \frac{\partial f}{\partial y} = 3y^2 - 5\cos(3x + 5y).$$

Higher order derivatives are defined by iterating this process. For example, for this example we have

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(2x - 3\cos(3x+5y)\right) = 2 + 9\sin(3x+5y) \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(3y^2 - 5\cos(3x+5y)\right) = 15\sin(3x+5y) \\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(3y^2 - 5\cos(3x+5y)\right) = 6y + 25\sin(3x+5y) \\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(2x - 3\cos(3x+5y)\right) = 15\sin(3x+5y)
\end{aligned}$$

In general, working with partial derivatives requires slightly stronger assumptions than in the corresponding single variable situations, and it is simplest to formulate these extra assumptions in terms of continuity of the partial derivatives. We say a function is a C^1 function if it has first order partial derivatives with respect to all its variables, and these partials are *continuous*. We say a function is a C^2 function if it has continuous first and second order partials, and so on.

Also, working with partial derivatives, especially if the function f is a vector function, leads to very cumbersome notation. We can often simplify this by using the *matrix derivative*, also known as the *Jacobian matrix*: Suppose f is a function of n variables and its values are m-dimensional vectors, which we regard as **column vectors**. If we calculate the partial derivative of f with respect to one of its variables we will get a m dimensional vector. We can then form the $m \times n$ matrix that has these partial derivatives as its columns. This matrix is written as Df, or as $Df(a)$ if the derivatives are all taken at the point a.

For example, if $f(x,y) = \begin{bmatrix} x^2 \\ xy + y^2 \end{bmatrix}$ then

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 2x \\ y \end{bmatrix},\ \frac{\partial f}{\partial y} = \begin{bmatrix} 0 \\ x + 2y \end{bmatrix},\ Df = \begin{bmatrix} 2x & 0 \\ y & x+2y \end{bmatrix},\ Df(-1,3) = \begin{bmatrix} -2 & 0 \\ 3 & 5 \end{bmatrix}.$$

Here are some of the main facts about partial derivatives:

THEOREM A.3. *If f is C^1 then f is continuous; if f is C^2 then f is also C^1.*

THEOREM A.4 (Equality of mixed partials). *If f is C^2 then the "mixed partials" with respect to any two varialbes are equal. That is,* $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$.

THEOREM A.5 (Chain rule). *If f and g are C^1 functions then the composition $f \circ g$ is C^1. If a is in the domain of g and $g(a)$ is in the domain of f then $D(f \circ g)(a) = Df(g(a))Dg(a)$.*

THEOREM A.6 (Linear approximation). *If f is C^1 and a is in the domain of f then*

$$f(x) = f(a) + Df(a)(x-a) + R(x)$$

where the remainder term $R(x)$ satisfies

$$\|R(x)\| = \varepsilon(x)\,\|x-a\|\,, \quad \lim_{x\to a} \varepsilon(x) = 0.$$

The condition on the remainder means that $R(x)$ is very small in comparison to $x-a$, if x is near a.

A.3. Exact differential equations

Here is a technique for finding a constant of the motion for a system of equations of the form $\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = f(x,y)$. We are going to treat this as an "exact differential equation". A similar method is covered in Calc III, or in Physics courses, in terms of "finding a potential function for a conservative force field".

Obtain an equation for $\frac{dy}{dx}$, and write it in the form $P(x,y)\,dx + Q(x,y)\,dy = 0$; this equation is satisfied along the trajectories $(x(t), y(t))$ of the dynamical system. A constant of the motion φ will satisfy $d\varphi = \frac{\partial\varphi}{\partial x}\,dx + \frac{\partial\varphi}{\partial y}\,dy = 0$ along these trajectories. This suggests that we try to find a function φ satisfying $d\varphi = P\,dx + Q\,dy$. If this is possible then $\frac{\partial\varphi}{\partial x} = P$ and $\frac{\partial\varphi}{\partial y} = Q$, so

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial\varphi}{\partial x}\right) = \frac{\partial^2\varphi}{\partial y\partial x} = \frac{\partial^2\varphi}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial\varphi}{\partial y}\right) = \frac{\partial Q}{\partial x}.$$

Hence this method can't work unless the *integrability condition* $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ is satisfied. Check this, and don't bother to continue unless it is correct. However, you can try to algebraically modify the equation $P(x,y)\,dx + Q(x,y)\,dy = 0$ to a different form $P_1(x,y)\,dx + Q_1(x,y)\,dy = 0$, and then you can try again. There are infinitely many ways to rewrite the equation, since you can multiply by any function of (x,y).

Now suppose the integrability condition is satisfied. Start with one of the equations, say $\frac{\partial\varphi}{\partial x} = P(x,y)$, and integrate **with respect to** x. The result will be $\varphi(x,y)$, involving an arbitrary constant **with respect to** x. That is, the arbitrary constant will in fact be a function of y, say $g(y)$. Now plug this expression for φ, including the unknown $g(y)$, into $\frac{d\varphi}{dy} = Q(x,y)$. All x terms should cancel out from this equation, leaving a differential equation for $g(y)$ in terms of y. (If not, you made a mistake; try again.) Solve this differential equation for $g(y)$; you don't need to keep the constant of integration. Since you know $c(y)$ you have now determined $\varphi(x, y)$. Congratulations: This is a constant of the motion.

Here's an example: Suppose $\dfrac{dx}{dt} = 2x - y^2$, $\dfrac{dy}{dt} = x - 2y$. Then

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{x - 2y}{2x - y^2}.$$

Multiplying this out and gathering all terms on the left hand side produces

$$(*) \qquad (2y - x)\,dx + (2x - y^2)\,dy = 0,$$

so the equations for φ are $\dfrac{\partial\varphi}{\partial x} = P(x, y) = 2y - x$, $\dfrac{\partial\varphi}{\partial y} = Q(x, y) = 2x - y^2$. The integrability condition $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$ is satisfied; both partials are equal to 2. If we integrate $\dfrac{\partial\varphi}{\partial x} = 2y - x$ with respect to x we find

$$\varphi = \int (2y - x)\,dx = 2xy - \tfrac{1}{2}x^2 + g(y)$$

where $g(y)$, an unknown function of y, is a constant as long as we consider x to be the variable. Now if we plug this expression for φ into the other equation, $\dfrac{\partial\varphi}{\partial y} = Q(x, y) = 2x - y^2$, we get

$$\frac{\partial\varphi}{\partial y} = \frac{\partial}{\partial y}\big(2xy - \tfrac{1}{2}x^2 + g(y)\big) = 2x + g'(y) = Q(x, y) = 2x - y^2.$$

This simplifies to $g'(y) = -y^2$, which integrates to $g(y) = \frac{1}{3}y^3$ (setting the constant of integration to 0). So

$$\varphi(x, y) = 2xy - \tfrac{1}{2}x^2 + g(y) = 2x - \tfrac{1}{2}x^2 - \tfrac{1}{3}y^3.$$

Notice that we were fortunate in the way we wrote the differential equation $(*)$. We might have written it as

$$dy - \frac{x - 2y}{2x - y^2}\,dx = 0, \quad \text{or} \quad \frac{dx}{2x - y^2} - \frac{dy}{x - 2y} = 0,$$

and neither of these satisfies the integrability condition.