ARTICLES

Hammer Juggling, Rotational Instability, and Eigenvalues

CARL V. LUTZER
Rochester Institute of Technology
Rochester, NY 14623
Carl.Lutzer@rit.edu

Introduction

Get a hammer. Seriously, get a hammer. As an experiment, hold the hammer in front of you with its head pointing up. Toss it upward (CAREFULLY!), end-over-end, and catch it after one revolution. The orientation of the hammer when you catch it will be the same as when you tossed it.

As a second experiment, hold the hammer in front of you with its head pointing sideways, to the right. Toss the hammer upward, end-over-end, and catch it after one revolution. This time, the orientation changes—the head pointed to the right when you tossed it, but points to the left when you catch it!

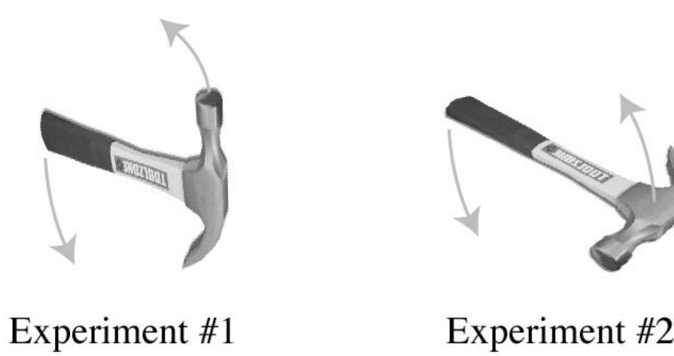


Figure 1 Hammer juggling and unstable rotation

Many people suggest that this strange 1/2-twist in experiment #2 is due to the asymmetry of the hammer's mass distribution, but the same kind of thing will happen with a book, or wallet, or any object with three distinct dimensions. (Try it! Use a rubber-band to keep the wallet or book closed.) We don't always see a *half*-twist (that will depend on the particular orientation of the object when you release it), but we almost always see a twist. Why? The answer is well known to the physics community, but is documented primarily in their parlance. The following exposition explains this phenomenon from a mathematician's point of view. The governing equations will be quickly derived, and the supporting linear algebra will be explored.

We assume that the reader has basic knowledge of multivariate calculus, and is aware that $e^{i\phi} = \cos\phi + i\sin\phi$. We also assume that the reader is familiar with eigenvalues, eigenvectors, linear independence, and understands that a proper choice of basis will diagonalize a symmetric matrix $M \in \mathbb{R}^{3\times 3}$.

The basics

In this section we begin with simple definitions of basic vocabulary, cite of the governing equations of motion, and then proceed with the salient calculations. Proofs of important assertions, and a derivation of the equations of motion are postponed until later sections so that we can focus on answering the question of why the hammer performs a half-revolution in Experiment #2 but not in Experiment #1.

Vocabulary

Angular Velocity Suppose an object is revolving about some particular axis, much like a child's spinning top. The *angular velocity* of the object, denoted by ω, is a vector that points in the direction of that axis. The magnitude of ω is $2\pi\gamma$, where $\gamma \geq 0$ is the number of revolutions per second. As you might infer from the example of the spinning top, the angular velocity vector may change direction and length as time evolves.

Newton's Second Law Most people cite Newton's Second Law as $F = ma$, which isn't quite right. Newton's Second Law says that force is the instantaneous change in momentum. In the case of linear force we write $F = d\rho/dt$ where $\rho = mv$ is the linear momentum of a mass m traveling with velocity v. In the case of angular force and angular momentum we write $\tau = dL/dt$ where τ means torque and L denotes angular momentum (discussed in detail later).

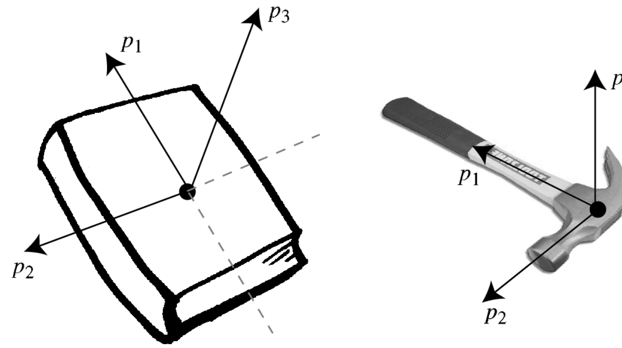
Euler's equation For reasons that will be explained later, the governing equation of motion is

$$\tau = M\dot{\omega} + \omega \times M\omega, \tag{1}$$

where $M \in \mathbb{R}^{3\times 3}$ is a symmetric matrix and $\dot{\omega}$ denotes the derivative of ω with respect to time. (This "dot notation" is used throughout the rest of the article to denote differentiation with respect to time.) In later sections we'll see that (1), called *Euler's equation*, is just a fancy restatement of the fact that $\tau = dL/dt$.

Calculations Because the matrix M is symmetric, its eigenvalues are all real, and eigenvectors associated with distinct eigenvalues are orthogonal. In fact, it happens that all the eigenvalues of M are positive! In the case of the hammer, they're also distinct so we label them in increasing order: $0 < \lambda_1 < \lambda_2 < \lambda_3$.

Physicists refer to M as the *moment-of-inertia tensor*, and they often use the letter I (for "inertia") to denote this matrix. (We use M in this exposition to avoid confusion with the identity matrix.) The eigenvalues of M are called the *principal moments of inertia,* and their corresponding unit-eigenvectors are called the *principal axes of rotation.* These unit-eigenvectors, which we'll denote by p_1, p_2, and p_3 respectively, point along "the axes of" the object in question. For example, pull a textbook off of the shelf. It has length, width, and height. The vector p_1 points in the direction of the length, the vector p_2 points in the direction of the width, and the vector p_3 points in the direction of the height (see the figure, below). Notice that, listed in the order prescribed by our indexing, the dimensions of the book are decreasing: length > width > height. If you accepted the earlier invitation to try the experiment with another object (with three distinct dimensions), you found that the rotation was unstable when the axis of rotation was parallel to p_2, which corresponds to the "middle" dimension. This will always be the case, as we'll see in a moment.



Vectors p_1, p_2, p_3 form an orthonormal basis for $\mathbb{R}^3$, so any angular velocity can be expressed as a linear combination of them: $\omega = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$. (Recall that ω may change with time, so the scalars α_1, α_2 and α_3 are functions of time.) Moreover, the matrix M is diagonal in the basis $\{p_1, p_2, p_3\}$.

$$M = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

So when the rotation is free from external torque and we use $\{p_1, p_2, p_3\}$ as our basis, equation (1) becomes

$$\lambda_1 \dot{\alpha}_1 + (\lambda_3 - \lambda_2)\alpha_2\alpha_3 = 0 \tag{2}$$

$$\lambda_2 \dot{\alpha}_2 + (\lambda_1 - \lambda_3)\alpha_1\alpha_3 = 0 \tag{3}$$

$$\lambda_3 \dot{\alpha}_3 + (\lambda_2 - \lambda_1)\alpha_1\alpha_2 = 0 \tag{4}$$

Suppose the object in question (the hammer, in this case) were to rotate about the axis p_1. Then $\alpha_2(0) = 0 = \alpha_3(0)$ and it follows from equations (2)–(4) that α_2 and α_3 *stay* zero. Of course, we see the same behavior whether we rotate about p_1, p_2 or p_3. But rotating about one of the principal axes—*exactly*—is highly unlikely, even if we are meticulous in our efforts to make it happen. So what happens when the object in question rotates about an axis that is very *close* to one of the principal axes?

Stable rotation Suppose ω is initially very close to p_1. Then $\alpha_2(0) \approx \alpha_3(0) \approx \varepsilon \approx 0$, so the second summand on the right-hand side of (2) is order ε^2.

$$\lambda_1 \dot{\alpha}_1 + \underbrace{(\lambda_3 - \lambda_2)\alpha_3\alpha_2}_{\mathrm{O}(\varepsilon^2)} = 0. \tag{5}$$

The analogous terms in (3) and (4) are only order ε, so a linear approximation of Euler's equation is

$$\lambda_1 \dot{\alpha}_1 \approx 0 \tag{6}$$

$$\lambda_2 \dot{\alpha}_2 + (\lambda_1 - \lambda_3)\alpha_1\alpha_3 = 0 \tag{7}$$

$$\lambda_3 \dot{\alpha}_3 + (\lambda_2 - \lambda_1)\alpha_1\alpha_2 = 0 \tag{8}$$

Equation (6) indicates that α_1 is constant (or nearly so). This reduces the problem to a system of two equations in two unknowns. Solving (7) and (8) for $\dot{\alpha}_2$ and $\dot{\alpha}_3$, respectively, gives us

$$\begin{pmatrix} \dot{\alpha}_2 \\ \dot{\alpha}_3 \end{pmatrix} = \begin{bmatrix} 0 & \dfrac{(\lambda_3 - \lambda_1)\alpha_1}{\lambda_2} \\ \dfrac{(\lambda_1 - \lambda_2)\alpha_1}{\lambda_3} & 0 \end{bmatrix} \begin{pmatrix} \alpha_2 \\ \alpha_3 \end{pmatrix} \tag{9}$$

which we write as the 2×2 system $\dot{x} = Ax$. The eigenvalues of A are

$$\pm i\sqrt{\frac{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)\alpha_1^2}{\lambda_2\lambda_3}},$$

which we will denote by $\pm i\phi$. Suppose the associated eigenvectors are $\vec{a}_1, \vec{a}_2 \in \mathbb{C}^2$. Then, since these vectors are linearly independent, there are scalars $c_1, c_2 \in \mathbb{C}$ such that $c_1\vec{a}_1 + c_2\vec{a}_2 = (\alpha_2(0), \alpha_3(0))^T$. Note that c_1 and c_2 are "small" since $\|\vec{a}_1\| = \|\vec{a}_2\| = |e^{\pm i\phi}| = 1$ and $\alpha_2(0) \approx 0 \approx \alpha_3(0)$. Now by defining $x(t) = c_1 e^{i\phi t}\vec{a}_1 + c_2 e^{-i\phi t}\vec{a}_2$ we have

$$\begin{aligned} \dot{x}(t) &= \frac{d}{dt}\left(c_1 e^{i\phi t}\vec{a}_1 + c_2 e^{-i\phi t}\vec{a}_2\right) = c_1\left(\frac{d}{dt}e^{i\phi t}\right)\vec{a}_1 + c_2\left(\frac{d}{dt}e^{-i\phi t}\right)\vec{a}_2 \\ &= c_1 e^{i\phi t}(i\phi)\vec{a}_1 + c_2 e^{-i\phi t}(-i\phi)\vec{a}_2 = c_1 e^{i\phi t}A\vec{a}_1 + c_2 e^{-i\phi t}A\vec{a}_2 \\ &= A(c_1 e^{i\phi t}\vec{a}_1 + c_2 e^{-i\phi t}\vec{a}_2) = Ax(t) \end{aligned}$$

The function $x(t)$ solves (9) with the correct initial data so, since that solution is unique, $x(t) = (\alpha_2(t), \alpha_3(t))^T$. It follows that α_2 and α_3 not only start small but *stay* small. That is, ω *stays* close to $\alpha_1 p_1$.

In fact, ω revolves around p_1 as the system evolves. It's easy to follow through the same calculations to derive the same behavior when the axis of rotation is close to p_3, but something very different happens when ω is initially near p_2.

Unstable rotation If we begin with ω very near to p_2, $\alpha_1(0) \approx 0 \approx \alpha_3(0)$, so a linear approximation of Euler's equation is

$$\lambda_1\dot{\alpha}_1 + (\lambda_3 - \lambda_2)\alpha_2\alpha_3 = 0 \tag{10}$$

$$\lambda_2\dot{\alpha}_2 \approx 0 \tag{11}$$

$$\lambda_3\dot{\alpha}_3 + (\lambda_2 - \lambda_1)\alpha_1\alpha_2 = 0. \tag{12}$$

Equation (11) indicates that α_2 is constant (or nearly so). This reduces the problem to a system of two equations and two unknowns.

$$\begin{pmatrix} \dot{\alpha}_1 \\ \dot{\alpha}_3 \end{pmatrix} = \begin{bmatrix} 0 & \dfrac{(\lambda_2 - \lambda_3)\alpha_2}{\lambda_1} \\ \dfrac{(\lambda_1 - \lambda_2)\alpha_2}{\lambda_3} & 0 \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_3 \end{pmatrix} \tag{13}$$

The coefficient matrix has eigenvalues

$$\pm\sqrt{\frac{(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_2)\alpha_2^2}{\lambda_1\lambda_3}},$$

which we denote by $\pm\phi$. Suppose the associated eigenvectors are $\vec{a}_1, \vec{a}_2 \in \mathbb{R}^2$. Then the solution to (13) is $x = c_1 e^{\phi t}\vec{a}_1 + c_2 e^{-\phi t}\vec{a}_2$, where c_1 and c_2 are chosen to achieve $x(0) = (\alpha_1(0), \alpha_3(0))^T$. It's important to note that $c_2 e^{-\phi t}\vec{a}_2$ vanishes quickly but that

$c_1 e^{\phi t}\vec{a}_1$ grows exponentially. That is, though α_1 and α_3 started small, they don't stay that way, and it's exactly this instability that makes the hammer change its orientation.

Rolling up our sleeves

Now we undertake the task of supporting the assertions made about the matrix M (that it's symmetric and that all its eigenvalues are positive) and explaining Euler's equation. We begin by defining angular momentum and establishing its relationship to angular velocity.

The relationship between L and ω Suppose a rigid body rotates about the line through its center-of-gravity defined by the vector ω. Taking the center-of-gravity as our origin, an atom at $r_j = (x_j, y_j, z_j)$ has a linear velocity of $v_j = \omega \times r_j$ (see Figure 2). The *angular momentum* of that atom is defined to be $L_j = r_j \times m_j v_j$, where m_j is its mass. That is, $L_j = m_j(r_j \times (\omega \times r_j))$. Grinding through the cross products brings us to

$$L_j = \begin{bmatrix} m_j(y_j^2 + z_j^2) & -m_j x_j y_j & -m_j x_j z_j \\ -m_j x_j y_j & m_j(x_j^2 + z_j^2) & -m_j y_j z_j \\ -m_j x_j z_j & -m_j y_j z_j & m_j(x_j^2 + y_j^2) \end{bmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}. \qquad (14)$$

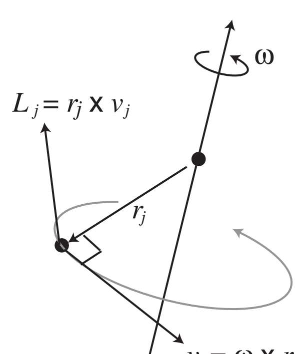


Figure 2 Angular velocity and angular momentum

The angular momentum of the entire object is just the sum of the angular momenta of all its atoms. Summing (14) over all particles gives us

$$L = \underbrace{\begin{bmatrix} \sum_j m_j(y_j^2 + z_j^2) & -\sum_j m_j x_j y_j & -\sum_j m_j x_j z_j \\ -\sum_j m_j x_j y_j & \sum_j m_j(x_j^2 + z_j^2) & -\sum_j m_j y_j z_j \\ -\sum_j m_j x_j z_j & -\sum_j m_j y_j z_j & \sum_j m_j(x_j^2 + y_j^2) \end{bmatrix}}_{\text{This is the matrix } M} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}.$$

Defining M to be the coefficient matrix on the right-hand side, we can write $L = M\omega$. We remark that the symmetry of M is now apparent, but why are its eigenvalues always positive and why does it play a role in Euler's equation? These questions are answered in the remaining sections.

The eigenvalues of M We begin our investigation into the eigenvalues of M by writing

$$M = (\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2)I - A^T A, \tag{15}$$

where $\vec{x}_j = \sqrt{m_j}\, x_j$, $\vec{y}$ and $\vec{z}$ are the corresponding vectors of scaled y and z coordinates, and A is the matrix whose columns are $A_{\cdot 1} = \vec{x}$, $A_{\cdot 2} = \vec{y}$, and $A_{\cdot 3} = \vec{z}$. That is, M is a perturbation of the matrix $(\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2)I$, which has a single eigenvalue whose algebraic multiplicity is three. The effect of this perturbation on the set of eigenvalues depends on the "size" of the perturbation. We measure the "size" of a linear function $\mathcal{L} : \mathbb{R}^3 \to \mathbb{R}^3$ with the *operator norm:*

$$\|\mathcal{L}\|_* \stackrel{\text{def}}{=} \max_{\|u\|=1} \|\mathcal{L}u\|, \tag{16}$$

where $\|v\| = \sqrt{v \cdot v}$ is the standard norm $\mathbb{R}^3$. (The fact that a maximum is always achieved follows from the Heine-Borel Theorem, which is usually taught in a course such as Real Analysis. Its 1-dimensional version is known to calculus students as the Extreme Value Theorem: *A continuous function on a closed interval achieves an absolute maximum value.*) Before continuing, we suggest that the reader verify the following lemma.

LEMMA 1. *Suppose $A, B : \mathbb{R}^3 \to \mathbb{R}^3$ are linear operators. Then*

1. $\|Ax\| \le \|A\|_* \|x\|$
2. $\|AB\|_* \le \|A\|_* \|B\|_*$
3. $\|A\|_* = \|A^T\|_*$

Now let us suppose that u is a unit-eigenvector of M associated with the eigenvalue λ. Then

$$\lambda u = Mu = \left((\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2)I - A^T A\right) u$$

from which it follows that $A^T A u = \left(\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2 - \lambda\right) u$. That is, u is an eigenvector of $A^T A$. The strategy of our proof is to use this fact to show that

$$\underbrace{\left| \underbrace{\left(\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2\right)}_{\text{anchor value} > 0} - \lambda \right|}_{\text{distance from } \lambda \text{ to anchor value}} < \|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2,$$

from which it follows that $\lambda > 0$. For example, if it was the case that $\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2 = 5$, showing $|5 - \lambda| < 5$ would imply that $\lambda > 0$.

Since $\|u\| = 1$, we have

$$\begin{aligned} \left| \|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2 - \lambda \right| &= \left\| \left(\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2 - \lambda\right) u \right\| \\ &= \|A^T A u\| \le \|A^T A\|_* \\ &\le \|A^T\|_* \|A\|_* = \|A^T\|_*^2 \end{aligned} \tag{17}$$

so the proof rests on our estimate of $\|A^T\|_*$. For any unit vector, v,

$$\|A^T v\| = \sqrt{(\vec{x}\cdot v)^2 + (\vec{y}\cdot v)^2 + (\vec{z}\cdot v)^2}$$
$$\leq \sqrt{\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2}. \tag{18}$$

Note that equality could only occur in (18) if some unit vector v were parallel (or antiparallel) to all three vectors, $\vec{x}$, $\vec{y}$ and $\vec{z}$. But this could only happen if the object in question were 1-dimensional! Restricting ourselves to 3-dimensional objects, we can rewrite (18) as

$$\|A^T v\| < \sqrt{\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2}. \tag{19}$$

Since (19) is true for all unit vectors v, it's true when $\|A^T v\|$ achieves its maximum and, thus, $\|A^T\|_* < \sqrt{\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2}$. Returning to (17), we have

$$\left|\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2 - \lambda\right| \leq \|A^T\|_*^2 < \|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2,$$

from which it follows that $\lambda \in (0, 2(\|\vec{x}\|^2 + \|\vec{y}\|^2 + \|\vec{z}\|^2)]$. That is, the eigenvalues of M are positive.

Euler's equation (explained) The final piece of the puzzle is Euler's equation which, earlier, we asserted was just a fancy way of saying that torque changes angular momentum. When we first introduced the idea of torque we wrote

$$\tau = \frac{dL}{dt}. \tag{20}$$

Equation (20) is correct from the point of view of an observer who is removed from the application of torque and the resulting change in motion—physicists say that such a person is in an *inertial frame*. But we're not dealing with an inertial frame because our coordinate system, $\{p_1, p_2, p_3\}$, depends on M, which depends on the object which is rotating. As the object rotates, so does our basis!

How do we write (20) from our point of view, at the center of the rotating body, with a basis that's rotating? The key is to imagine what an observer in an inertial frame would see if, from our point of view in the rotating basis, we saw no change in the angular momentum. Because our frame of reference is spinning, our observation that L appears to be constant means that L is spinning about the axis of revolution at exactly the same speed as the basis. So an observer in an inertial frame would record the change in angular momentum as $\omega \times L$ (see Figure 3). Using the subscript of 0 to denote the inertial frame and the subscript r to denote the rotating frame, this thought experiement allows us to write (20) from our point of view in the rotating frame:

$$\tau = \left(\frac{dL}{dt}\right)_0 = \underbrace{\left(\frac{dL}{dt}\right)_r}_{\text{our basis}} + \underbrace{\omega \times L_r}_{\text{is spinning}}. \tag{21}$$

Finally, we use the fact that $L_r = M\omega$. Notice that M depends on the physical characteristics of the object but not on time, so we can rewrite (21) as

$$\tau = M\dot{\omega} + \omega \times M\omega, \tag{22}$$

which is Euler's equation.

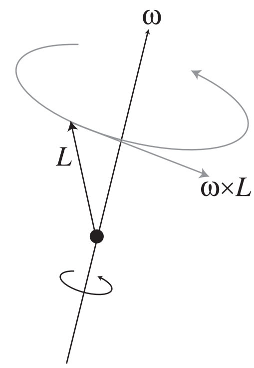


Figure 3 Change of L in an inertial frame

Conclusion

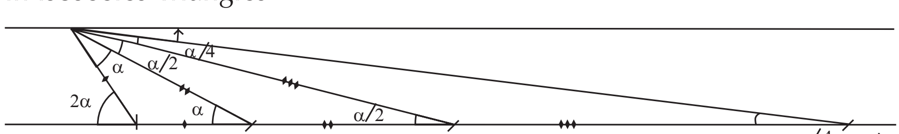
The last implicit supposition in our analysis was that the eigenvalues were distinct. This, at least, is not always true. What would happen if two of the eigenvalues were the same? What if all *three* were the same? What would that imply about the rotating object?

Acknowledgments. The author would like to thank Drs. Scott Franklin and George Thurston of RIT's Department of Physics for their time and conversation regarding this article.

REFERENCES

1. C.H. Edwards & D.E. Penney, *Elementary Differential Equations*, 5e, Prentice Hall, Upper Saddle River, NJ, 2004.
2. D.C. Lay, *Linear Algebra and Its Applications*, 3e, Addison Wesley, Boston, 2003.
3. D. Halliday, R. Resnick, J. Walker, *Fundamentals of Physics*, 6e, John Wiley & Sons, Inc., 2001.

Proof Without Words: Sum of a Geometric Series via Equal Base Angles in Isosceles Triangles



$$\alpha + \frac{\alpha}{2} + \frac{\alpha}{4} + \cdots = \sum_{n=0}^{\infty} \frac{\alpha}{2^n} = 2\alpha$$

—Ángel Plaza
ULPGC, 35017-Las Palmas G.C., Spain